VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS

Lech GULBINOVIČ

INFORMACINIŲ SISTEMŲ SAUGUMO TYRIMAS IR IŠLIEKAMUMO VERTINIMO MODELIO SUKŪRIMAS

DAKTARO DISERTACIJA

TECHNOLOGIJOS MOKSLAI,
INFORMATIKOS INŽINERIJA (07T)

VGTU
LEIDYKLA
Vilnius TECHNIKA 2014

Disertacija rengta 2010–2014 metais Vilniaus Gedimino technikos universitete.

Mokslinis vadovas

prof. habil. dr. Julius SKUDUTIS (Vilniaus Gedimino technikos universitetas, informatikos inžinerija – 07T).

Vilniaus Gedimino technikos universiteto Informatikos inžinerijos mokslo krypties disertacijos gynimo taryba:

Pirmininkas

prof. habil. dr. Antanas ČENYS (Vilniaus Gedimino technikos universitetas, informatikos inžinerija – 07T).

Nariai:

prof. dr. Algirdas BAŠKYS (Vilniaus Gedimino technikos universitetas, elektros ir elektronikos inžinerija – 01T),
prof. habil. dr. Gintautas DZEMYDA (Vilniaus universitetas, informatikos inžinerija – 07T),
doc. dr. Raimundas MATULEVIČIUS (Tartu universitetas, informatikos inžinerija – 07T),
prof. dr. Olegas VASILECAS (Vilniaus Gedimino technikos universitetas, informatikos inžinerija – 07T).

Disertacija bus ginama viešame informatikos inžinerijos mokslo krypties disertacijos gynimo tarybos posėdyje **2014 m. gruodžio 15 d. 14 val.** Vilniaus Gedimino technikos universiteto senato posėdžių salėje.

Adresas: Saulėtekio al. 11, LT-10223 Vilnius, Lietuva.
Tel.: (8 5) 274 4956; faksas (8 5) 270 0112; el. paštas doktor@vgtu.lt

Pranešimai apie numatomą ginti disertaciją išsiųsti 2014 m. lapkričio 14 d.

Disertaciją galima peržiūrėti interneto svetainėje http://dspace.vgtu.lt/ ir Vilniaus Gedimino technikos universiteto bibliotekoje (Saulėtekio al. 14, LT-10223 Vilnius, Lietuva).
VGTU leidyklos TECHNIKA 2285-M mokslo literatūros knyga
http://leidykla.vgtu.lt

ISBN 978-609-457-733-8

lech.gulbinovic@vgtu.lt

VILNIUS GEDIMINAS TECHNICAL UNIVERSITY

Lech GULBINOVIČ

INFORMATION SYSTEM SECURITY EVALUATION AND CREATION OF SURVIVABILITY EVALUATION MODEL

DOCTORAL DISSERTATION

TECHNOLOGICAL SCIENCES,
INFORMATICS ENGINEERING (07T)

VGTU
LEIDYKLA
Vilnius TECHNIKA 2014

Doctoral dissertation was prepared at Vilnius Gediminas Technical University in 2010–2014.

Scientific Supervisor

Prof Dr Habil Julius SKUDUTIS (Vilnius Gediminas Technical University, Informatics Engineering – 07T).

The Dissertation Defense Council of Sciencific Field of Informatics Engineering of Vilnius Gediminas Technical University:

Chairman

Prof Dr Habil Antanas ČENYS (Vilnius Gediminas Technical University, Informatics Engineering – 07T).

Members:

Prof Dr Algirdas BAŠKYS (Vilnius Gediminas Technical University, Electrical and Electronic Engineering – 01T),

Prof Dr Habil Gintautas DZEMYDA (Vilnius University, Informatics Engineering – 07T),

Assoc Prof Dr Raimundas MATULEVIČIUS (University of Tartu, Informatics Engineering – 07T),

Prof Dr Olegas VASILECAS (Vilnius Gediminas Technical University, Informatics Engineering – 07T).

The dissertation will be defended at the public meeting of the Dissertation Defense Council of Informatics Engineering in the Senate Hall of Vilnius Gediminas Technical University at **2 p. m. on 15 December 2014**.

Address: Saulėtekio al. 11, LT-10223 Vilnius, Lithuania.
Tel.: +370 5 274 4956; fax +370 5 270 0112; e-mail: doktor@vgtu.lt

A notification on the intend defending of the dissertation was send on 14 November 2014.

A copy of the doctoral dissertation is available for review at the Internet website http://dspace.vgtu.lt/ and at the Library of Vilnius Gediminas Technical University (Saulėtekio al. 14, LT-10223 Vilnius, Lithuania).

Reziumė

Disertacijoje nagrinėjamos informacinių sistemų saugumo problemos. Pagrindinis tyrimo objektas – informacinės sistemos išliekamumas ir jo įvertinimo priemonės. Darbo tikslui pasiekti – sukurti informacinės sistemos saugumo lygio įvertinimo modelį, yra svarbūs šie aspektai: saugumo mechanizmai, informacinių sistemų saugumo modeliavimo metodai, sistemų išliekamumas ir duomenų paketų pasiskirstymas tinkle. Darbe sprendžiami keli pagrindiniai uždaviniai: informacinės sistemos saugumo modeliavimo metodų analizė, kompiuterių tinklo srautų statistinė analizė, informacinės sistemos išliekamumo vertinimo modelio sukūrimas, tikros informacinės sistemos saugumo lygio įvertinimo charakteristikų gavimas.

Disertaciją sudaro įvadas, keturi skyriai, rezultatų apibendrinimas, naudotos literatūros ir autoriaus publikacijų disertacijos tema sąrašai.

Įvadiniame skyriuje aptariama tiriamoji problema, darbo aktualumas, aprašomas tyrimų objektas, formuluojamas darbo tikslas ir uždaviniai, aprašoma tyrimų metodika, darbo mokslinis naujumas, darbo rezultatų praktinė reikšmė, ginamieji teiginiai. Įvado pabaigoje pristatomos autoriaus paskelbtos publikacijos ir pranešimai nacionalinėse ir tarptautinėse konferencijose disertacijos tema bei disertacijos struktūra.

Pirmasis skyrius skirtas literatūros apžvalgai. Jame pateiktos kompiuterių tinklo apsaugos mechanizmų (priemonių), tinklo saugumo modeliavimo metodų ir tinklo statistikos panaudojimo analitinės apžvalgos. Skyriaus pabaigoje formuluojamos išvados ir tikslinami disertacijos uždaviniai.

Antrajame skyriuje pateiktas kompiuterių tinkle surinktos informacijos panaudojimo metodas statistinių duomenų apie tinklo srautus gavimui. Taip pat pateikti tinklo srautų statistinės analizės eksperimentiniai rezultatai. Eksperimentiniu būdu ir palyginimų metodu nustatytas paketų atėjimo laiko pasiskirstymo dėsnis.

Trečiajame skyriuje pateiktas informacinių sistemų saugumo modeliavimo modelis. Ištirtas ir patikrintas modelio veikimas. Pateiktas valstybinių institucijų informacinių sistemų kategorijų aprašymas ir palyginimas.

Ketvirtajame skyriuje pateikta metodika, kaip iš aptiktų incidentų statistikos gauti incidentų pasiskirstymo tikimybes įtakojančias saugumą. Pasiskirstymai pritaikyti naujajame modelyje. Gautos realios informacinės sistemos saugumo vertinimo charakteristikos.

Disertacijos tema paskelbti 6 straipsniai ir perskaityti 7 pranešimai nacionalinėse ir tarptautinėse konferencijose.

Abstract

The dissertation investigates the information security. The main objects of research include information system security mechanisms, information system modeling methods, information system survivability and packet distribution in computer network. These are the objects important to obtain the main aim and create a information security systems simulation tool for computer system security level evaluation.

The paper approaches a few major tasks: information security modeling method selection, computer network statistical analysis, the information system survivability evaluation model creation and real information system security level evaluation characteristics.

The dissertation consists: Introduction, 4 chapters, Conclusions and References.

The introduction reveals the investigated problem, importance of the thesis and the object of research and describes the purpose and tasks of the paper, research methodology, scientific novelty, the practical significance of results examined in the paper and defended statements. The introduction ends in presenting the author's publications on the subject of the defended dissertation, offering the material of made presentations in conferences and defining the structure of the dissertation.

Chapter 1 revises used literature. There are reviewed information system security mechanisms, network security modeling methods analysis and network statistic usage review. At the end of the chapter, conclusions are drawn and the tasks for the dissertation are reconsidered.

Chapter 2 provides the method, how collected information is used for statistics about network flows. There are provided computer network statistical analysis experiment results. From experiment the statistical packet time inter arrival distribution was determined.

Chapter 3 provides information system security modeling model. The model operation was checked. The government information system categories are described and they security evaluation were compared.

Chapter 4 provides a methodology how gather a statistical incident distribution from incident report. Distributions were used in model. The real information system security evaluation characteristics were obtained.

6 articles focusing on the subject of the discussed dissertation are published. 7 presentations on the subject have been given in conferences at national and international level.

Žymėjimai

Simboliai

α, β – grafiko formos parametrai;
α_{mi} – koeficientai įvertinantys modulio atsparumą incidentams;
A – daugiklis;
$A(t)$ – pasiekiamumo tikimybė;
A_{pf} – vidutinis paketų skaičius sraute;
A_{ps} – vidutinis paketų dydis sraute;
A_S – duomenų vidurkio pasiskirstymas;
b_n – sistemos būsena;
β_{Th} – koeficientai grėsmės tipo įvertinimui;
Δ_n – tarpas tarp paketų;
Δt – laiko periodas;
Δt_d – modulio sukompromitavimo aptikimo laiko intervalas;
$F(X)$ – pasiskirstymo funkcija;
$f(n)$ – paketų atėjimo laiko pasiskirstymas tarp paketų;
i – incidentai;
IAn – momentinė veikla;
IGn – įėjimo vartai;
j – sunkumo lygis;
M – aikštelė;
m – modulis;
n – saugumo mechanizmų kiekis;

N_F – srautų skaičius;
N_p – paketų skaičius;
$N_{m(n)}$ – saugumo mechanizmų rinkiniai;
OGn – išėjimo vartai;
P, λ, μ – tikimybės;
$P_M(m)$ – modulių panaudojimo tikimybės;
$P_m(j)$ – grėsmių sunkumas;
λ_{ij} – įvykių srauto intensyvumas;
S_i – sistemos būsena;
S – išliekamumas;
T – laikas;
TAn – laikinė veikla;
t_n – paketų atėjimo laikas;
$w(m)$ – modulio svarba;
X – atsitiktinis dydis.

Santrumpos

AN – veiklos tinklai (angl. *Activity Network*);
CDF – normuota pasiskirstymo funkcija (angl. *Cumulative Distribution Function*);
DoS – atsisakymo aptarnauti ataka (angl. *Denial of Service*);
DMZ – perimetro tinklas (angl. *Demilitarized Zone*);
FTP – failų perdavimo protokolas (angl. *File Transfer Protocol*);
GSPN – apibendrintas stochastinis Petri tinklas (angl. *Generalized Stochastic Petri Net*);
HTTP – protokolas internetinių puslapių perdavimui (angl. *Hypertext Transfer Protocol*);
ICMP – interneto valdymo žinučių protokolas (angl. *Internet Control Message Protocol*);
IDS – įsilaužimo aptikimo sistema (angl. *Intrusion Detection System*);
IP – interneto protokolas (angl. *Internet Protocol*);
KS – Kolmogorov-Smirnov parametras;
L3/L4 – trečias/ketvirtas atvirų sistemų sąveikos lygmuo;
OS – operacinė sistema (angl. *Operation System*);
OSI – atvirų sistemų sąveika (angl. *Open Systems Interconnection*);
PĮ – programinė įranga;

TCP	– transporto lygmens protokolas (angl. *Transport Control Protocol*);
TCP SYN	– užklausa sinchronizuoti sekos numerius;
SPN	– stochastiniai Petri tinklai (angl. *Stochastic Petri Nets*);
SAN	– stochastinės veiklos tinklai (angl. *Stochastic Activity Network*);
SNMP	– paprastas tinklo stebėjimo protokolas (angl. *Simple Network Management Protocol*);
UDP	– transporto lygmens protokolas (angl. *User Datagram Protocol*);
VGTU	– Vilniaus Gedimino technikos universitetas;
VPN	– virtualus privatus tinklas (angl. *Virtual Privat Network*).

Turinys

ĮVADAS 1
Problemos formulavimas 1
Darbo aktualumas 2
Tyrimų objektas 2
Darbo tikslas 2
Darbo uždaviniai 3
Tyrimų metodika 3
Darbo mokslinis naujumas 3
Darbo rezultatų praktinė reikšmė 4
Ginamieji teiginiai 4
Darbo rezultatų aprobavimas 4
Disertacijos struktūra 5
Padėka 5

1. INFORMACINIŲ SISTEMŲ APSAUGOS PRIEMONIŲ IR SAUGUMO MODELIAVIMO METODŲ ANALITINĖ APŽVALGA 7
1.1. Informacinių sistemų apsaugos priemonių ir incidentų analitinė apžvalga 7
1.1.1. Incidentai informacinėse sistemose 8
1.1.2. Apsaugos priemonės kompiuterių tinkle 10
1.2. Tinklo saugumo modeliavimo metodų analitinė apžvalga ir metodo pasirinkimas 13
1.2.1. Tikimybiniai metodai 14

1.2.2. Markovo procesai 16
1.2.3. Petri tinklai 17
1.2.4. Stochastinės veiklos tinklai 18
1.3. Tinklo srautų statistikos panaudojimas 20
1.4. Pirmojo skyriaus išvados ir disertacijos uždavinių formulavimas 22

2. KOMPIUTERIŲ TINKLO SRAUTŲ STATISTINĖ ANALIZĖ 23
2.1. Akademinis kompiuterių tinklas 24
2.2. Srautų kompiuterių tinkle statistinė analizė 27
2.3. Atėjimo laiko tarp paketų pasiskirstymo tyrimas 32
2.4. Pasiskirstymo matematinė išraiška 36
2.5. Antrojo skyriaus išvados 42

3. INFORMACINĖS SISTEMOS IŠLIEKAMUMO ĮVERTINIMAS REMIANTIS SAUGUMO ANALIZE 43
3.1. Saugumą reguliuojantis įstatymas 44
3.2. Informacinės sistemos modelio charakteristikos 45
3.3. Modelio sudarymas 46
3.4. Sistemos sukompromitavimo tikimybė 53
3.5. Trečiojo skyriaus išvados 59

4. AKADEMINIO TINKLO INFORMACINIŲ SISTEMŲ SAUGUMO LYGIO ĮVERTINIMAS 61
4.1. Informacinių sistemų patikimumas ir jo sudedamosios dalys 61
4.2. Saugumo incidentai tinkle 64
4.3. Informacinės sistemos išliekamumas 67
4.4. Ketvirtojo skyriaus išvados 74

BENDROSIOS IŠVADOS 77

LITERATŪRA IR ŠALTINIAI 79

AUTORIAUS MOKSLINIŲ PUBLIKACIJŲ DISERTACIJOS TEMA SĄRAŠAS 87

SUMMARY IN ENGLISH 89

PRIEDAI 107
A priedas. Informacinės sistemos saugumo incidentai. Lentelės ir grafikai 108
B priedas. Informacinės sistemos saugumo reikalavimai 121
C priedas. Bendraautorių sutikimai teikti publikacijų medžiagą disertacijoje 125
D priedas. Autoriaus mokslinių publikacijų disertacijos tema kopijos 139

Contents

INTRODUCTION 1
Problem formulation 1
Relevance of the thesis 2
Object of research 2
The aim of the thesis 2
Objectives of the thesis 3
Research methodology 3
Sciencific novelty of the thesis 3
Practical value of research findings 4
Defended statements 4
Approval of research findings 4
The structure of the thesis 5
Gratitude 5

1. INFORMATION SYSTEM SECURITY MODELING TECHNIQUES AND MODELING METHODS OVERVIEW 7
1.1. Informatic system security mechanisms and incedents analytical analysis 7
1.1.1. Incidents in informatic system 8
1.1.2. Security mechanisms in computer network 10
1.2. Network security moddeling methods analytical analysis 13
1.2.1. Probability methods 14
1.2.2. Markov process 16
1.2.3. Petri network 17

1.2.4. Stochastic activity network 18
1.3. Network flow statistic usage 20
1.4. Conclusions of the chapter 1 22

2. COMPUTER NETWORK TRAFFIC STATISTICAL ANALYSIS 23
2.1. Academic computer network 24
2.2. Network flow statistical analysis 27
2.3. Packet arrival time evaluation 32
2.4. Distribution mathematical expression 36
2.5. Conclusions of the chapter 2 42

3. INFORMATIONS SYSTEM SURVIVABILITY EVALUATION ACCORDING RISK ANALYSIS 43
3.1. Security regulation 44
3.2. Informatic system model characteristics 45
3.3. Creation of simulation model 46
3.4. Informatic system survivability 53
3.5. Conclusions of the chapter 3 59

4. ACADEMIC INFORMATION SYSTEM SECURITY EVALUATION 61
4.1. Informatic system dependability 61
4.2. Security incidents in network 64
4.3. Informatic system survivability 67
4.4. Conclusions of the chapter 4 74

GENERAL CONCLUSIONS 77

REFERENCES 79

LIST OF PUBLICATIONS BY THE AUTHOR ON THE TOPIC OF THE DISSERTATION 87

SUMMARY IN ENGLISH 89

ANNEXES 107
Annex A. Informaction system security incidents. Tables and figures 108
Annex B. Security requirements for information system 121
Annex C. The co-authors consents to present for the dissertation 125
Annex D. Copies of scientic publications by the author on the topic of the dissertation 139

Įvadas

Problemos formulavimas

Informacinės sistemos nuolat sudėtingėja, dėl naujų technologijų, protokolų, standartų ir augančio informacinių sistemų poreikio. Atsiranda nauji saugumą užtikrinantys metodai ir mechanizmai, pradedant nuo fizinio saugumo lygio užtikrinimo iki programinio lygio. Taip pat nuolat didėja saugumą įtakojančių veiksnių skaičius ir įvairovė. Yra didelis spektras galimybių saugumui užtikrinti. Tačiau yra sunku nustatyti kiek ir kokių saugumą užtikrinančių mechanizmų reikia, norint užtikrinti tam tikrą informacinės sistemos saugumo lygį.

Informacinė sistema – šiame darbe laikoma sistema, kuri sudaryta iš techninės įrangos, skirtos informacijos saugojimui, įvedimui ir išvedimui, programinės įrangos, skirtos informacijos įvedimui, ieškojimui, apdorojimui, išvedimui ir vaizdavimui bei telekomunikacijų tinklo, skirto informacijos perdavimui tarp sistemos ar jos dalių. Dažniausiai prieš projektuojant informacines sistemas rekomenduojama atlikti rizikos analizę, kurios pasėkoje nustatomi saugumo reikalavimai sistemai ir saugumą užtikrinantis mechanizmai. Tokia analizė reikalauja nemažai papildomo laiko ir kaštų, dėl to saugumo lygio tyrimai atliekami retai, kas įtakoja per mažą arba per didelį saugumo mechanizmų kiekį, dėl to pasiekiamas nepakankamas saugumo lygis arba per

sukuriama per daug sudėtinga ir brangi sistema. Reikalingas pakankamai greitas, pigus ir patogus būdas, leidžiantis dar informacinės sistemos projektavimo stadijoje nustatyti saugumo mechanizmų kiekį bei jų tipus, kad užtikrinti pageidaujamą sistemos saugumo lygį. Vienas iš problemos sprendimo būdų yra informacinės sistemos saugumo modelis, kurį galima naudoti dar sistemos projektavimo metu.

Disertacijoje, remiantis incidentų kompiuterių tinkluose ir apsaugos mechanizmų (priemonių) analize, sudarytas informacinės sistemos saugumo modelis, kuris gali būti pritaikomas projektuojant įvairias informacines sistemas.

Darbo aktualumas

Projektuojant, eksploatuojant ir prižiūrint informacines sistemas, informacijos saugumas bei duomenų išliekamumas yra laikomi aukščiausiu prioritetu. Reikalavimai privačiose bei valstybinėse organizacijose, kuriamai produkcijai bei paslaugoms informacijos saugumo ir patikimumo srityse yra gan aukšti. Tokiomis sistemomis gali būti: bankomatas, medicinos aparatas, lėktuvo arba branduolinio reaktoriaus kompiuterizuotos valdymo sistemos.

Dažniausiai tokios sistemos tarpusavyje yra sujungtos per kompiuterių tinklą, dėl to visos sistemos darbas yra labai priklausomas nuo tinklo saugumo ir patikimumo. Savo ruožtu toks platus kompiuterių tinklų panaudojimas iššaukia padidėjusį potencialių įsibrovėlių į tinklus susidomėjimą, dėl to tinklų pažeidžiamumas nuolat didėja.

Labai svarbu žinoti informacinių sistemų saugumo sudedamąsias dalis, incidentų pasiskirstymą tinkle, jų įtaką saugumui. Turint projektuojamos informacinės sistemos saugumo modelį pagreitėja kūrimo procesas, padidėja saugumas bei mažėja patiriami kaštai.

Tyrimų objektas

Darbo tyrimų objektas – informacinės sistemos išliekamumas ir jo įvertinimo priemonės.

Darbo tikslas

Darbo tikslas – sukurti įrankį, kuris leistų įvertinti informacinės sistemos išliekamumą dar projektavimo etape, remiantis saugumo reikalavimais bei įvertinant egzistuojančius saugumo mechanizmus ir incidentus.

Darbo uždaviniai

Darbo tikslui pasiekti, darbe reikia spręsti šiuos uždavinius:

1. Išanalizuoti tinkle egzistuojančius incidentus ir nustatyti jų įtaką saugumo sudedamosioms dalims – konfidencialumui, vientisumui ir pasiekiamumui.
2. Išanalizuoti informacinių sistemų saugumo modeliavimo metodus.
3. Atlikti kompiuterių tinklo duomenų srautų statistinę analizę, rasti TCP ir UDP paketų pasiskirstymo dėsnius.
4. Sudaryti informacinių sistemų išliekamumo įvertinimo modelį.
5. Nustatyti informacinės sistemos saugumo lygio vertinimo charakteristikas.

Tyrimų metodika

Darbe kiekybiniam informacinės sistemos saugumo lygio vertinimui taikomi tikimybiniai ir statistinės analizės metodai. Atsitiktiniams įvykiams informacinėse sistemose modeliuoti naudojami stochastinės veiklos tinklai (angl. *Stochastic Activity Networks – SAN*). SAN modeliai sudaryti panaudojant stochastinių veiklos tinklų modeliavimo įrankį *Mobius*. Surinktą informaciją apdorojus *NetFlows* protokolu, atlikta tinklo srautų statistinė analizė.

Darbo mokslinis naujumas

Rengiant disertaciją buvo gauti šie informatikos inžinerijos mokslui nauji rezultatai:

1. Sudarytas informacinės sistemos išliekamumo modeliavimo metodas, kuris leidžia nustatyti sistemos sukompromitavimo tikimybes remiantis grėsmės tipu, sunkumu ir saugumo mechanizmų rinkiniu.
2. Panaudotas *NetFlows* protokolas tinklo srauto statistikos surinkimui didelės apkrovos tinkluose, o surinkta informacija panaudota paketų pasiskirstymo analizei. Iš gautos statistikos nustatyti TCP ir UDP srauto paketų pasiskirstymo dėsniai akademiniame tinkle.
3. Pasiūlytos formulės, leidžiančios pritaikyti sistemos sukompromitavimo tikimybes remiantis grėsmės tipu, sunkumo ir saugumo mechanizmų

rinkiniu. Formulės gali būti lengvai keičiamos ir pritaikytos realios informacinės sistemos poreikiams.

Darbo rezultatų praktinė reikšmė

Sukurtas ir patikrintas informacinės sistemos išliekamumo modelis, kuris leidžia įvertinti sistemos išliekamumą priklausomai nuo šių parametrų: incidentų pasiskirstymo, incidentų dažnio, incidentų svorio, incidentų tipo, sistemos modulių svorio, apsaugos mechanizmų skaičiaus, incidentų aptikimo laiko, modulių atstatymo laiko.

Remiantis gauta informacinės sistemos išliekamumo charakteristika, galima nustatyti koks turi būti įgyvendintas saugumo mechanizmų kiekis sistemoje, kad ji tenkintų keliamus išliekamumo reikalavimus padidėjus incidentų skaičiui kompiuterių tinkle.

Ginamieji teiginiai

1. Pareto 2 pasiskirstymą galima naudoti TCP ir UDP tinklo srauto paketų atėjimo laiko dėsnio aprašymui akademiniame tinkle.
2. Pasiūlyta metodika gali būti naudojama informacinės sistemos išliekamumo įvertinimui, turint sistemos saugumo reikalavimus ir incidentų pasiskirstymą tinkle.
3. Informacinės sistemos saugumo lygis gali būti vertinamas pagal jos išliekamumo charakteristiką.

Darbo rezultatų aprobavimas

Disertacijos tema yra atspausdinti šeši moksliniai straipsniai: keturi – recenzuojamuose periodiniuose ir du kituose leidiniuose.

Disertacijoje atliktų tyrimų rezultatai buvo pristatyti septyniose mokslinėse konferencijose Lietuvoje ir užsienyje:

- Atakų atpažinimo sistemų analitinė apžvalga. 14-oji Lietuvos jaunųjų mokslininkų konferencija, Vilnius 2011.
- Network security analysis modeling techniques. Electronics 2011. The 15th International Conference of Electronics, Vilnius 2011.
- System Survivability Evaluation Based on Risk Analysis. Information Systems Architecture and Technology 2011, Szklarska Poręba 2011.

- Academic Computer Network Traffic Statistical Analysis. 2nd Baltic Congress on Future Internet Communications 2012, Vilnius 2012.
- Packet Inter-arrival Time Distribution in Academic Computer Network. 16-oji Lietuvos jaunųjų mokslininkų konferencija, Vilnius 2013.
- Packet Inter-arrival Time Distribution in Academic Computer Network. The 17th International Conference of Electronics, Palanga 2013.
- Akademinio tinklo kompiuterių sistemų saugumo lygio įvertinimas. 17-oji Lietuvos jaunųjų mokslininkų konferencija, Vilnius 2014.

Disertacijos struktūra

Disertaciją sudaro įvadas, keturi skyriai ir rezultatų apibendrinimas. Darbo pabaigoje pateiktas literatūros šaltinių ir autoriaus publikacijų disertacijos tema sąrašas.

Darbo apimtis yra 107 puslapiai neįskaitant priedų, tekste panaudotos 7 numeruotos formulės, 41 paveikslų ir 13 lentelių. Rašant disertaciją buvo panaudota 90 literatūros šaltinių.

Padėka

Dėkoju mokslinio darbo vadovui prof. habil. dr. Juliui Skudučiui ir doc. dr. Nerijui Paulauskui už konsultacijas, vertingas pastabas ir pagalbą rašant disertaciją.

Nuoširdžiai dėkoju darbo recenzentams doc. dr. Nikolajui Goraninui ir prof. habil. dr. Gintautui Dzemydai už vertingas pastabas, komentarus ir pasiūlymus. Taip pat nuoširdžiai dėkoju doc. dr. Eimantui Garšvai ir doc. dr. Gediminui Gražulevičiui už kryptingą pagalbą atliekant tyrimus.

Dėkoju savo šeimai – mamai ir tėvui – už tikėjimą ir palaikymą. Ypač dėkoju savo gyvenimo draugei Jolantai už kantrybę, supratingumą ir pagalbą galutinai užbaigiant disertacijos darbą.

1

Informacinių sistemų apsaugos priemonių ir saugumo modeliavimo metodų analitinė apžvalga

Šiame skyriuje yra atlikta informacinių sistemų apsaugos mechanizmų analitinė apžvalga. Aprašyti populiariausių atakų tipai, aptartos apsaugos priemonės nuo atakų. Išnagrinėti žinomi modeliavimo metodai, aptartos jų galimybės, privalumai bei trūkumai ir pasirinktas tinkamiausias pagrindinių saugumo veiksnių įvertinimo metodas. Skyriaus pabaigoje pateikiamos pirmojo skyriaus išvados ir formuluojami disertacijos uždaviniai. Skyriaus medžiaga publikuota autoriaus straipsnyje (Gulbinovič 2012).

1.1. Informacinių sistemų apsaugos priemonių ir incidentų analitinė apžvalga

Spartaus šiuolaikinių informacinių technologijų vystymosi dėka atskiri kompiuteriai ir lokalieji tinklai susijungė į vieningus korporacijų tinklus. Greta akivaizdžių privalumų toks susijungimas iššaukė eilę problemų, būdingų korporacijų tinklams. Viena svarbiausių problemų tapo informacijos saugumas.

Informacijos perdavimą, apdorojimą bei saugojimą reglamentuoja tam tikros taisyklės ir normos. Sistemos saugumo reikalavimus nusako įvairūs standartai (ISO 27001), taisyklės (ISP 2012), įstatymai, rekomendacijos (ISO 27002; ISO 27005; ISO 27033) ir kt. Tokių taisyklių, įstatymų ir praktinių rekomendacijų visuma, apimanti visus informacijos apdorojimo proceso aspektus, vadinama saugumo politika. Saugumo politika gali apimti tiek visą informacine sistemą, tiek atskiras jos sudedamąsias dalis (NCB 2011; ISP 2006;). Šių taisyklių ir normų laikymasis leidžia apsisaugoti nuo daugybės grėsmių ir užtikrina būtiną, o kartais net pakankamą sistemos saugumą.

Bet kurios informacinės sistemos saugumo politika yra formuojama jos projektavimo etape. Vėliau saugumo politika transformuojasi į apsaugos modelius, scenarijus, servisus bei mechanizmus, kurie realizuojami vėlesniuose projektavimo etapuose. Čia saugumo politikai yra priskiriama tiktai informacijos, jos resursų bei juos palaikančio tinklo infrastruktūros apsauga. Saugumo politiką (informacijos apdorojimo technologijų apsaugą) užtikrinančios priemonės yra tarp segmentiniai ekranai, atakų atpažinimo, srauto šifravimo sistemos, antivirusinės sistemos, fizinė apsauga, „mobilaus kodo“ kontrolės sistemos ir kt.

Dabartiniame tinkle daugiau dėmesio pradedama skirti saugumo didinimo priemonėms. Yra labai didelė saugumą užtikrinančių priemonių ir jų taikymo sričių įvairovė (Paulauskas *et al.* 2009). Egzistuoja labai daug įvairių besiskiriančių savo veikimu saugumą užtikrinančių sistemų ir metodų, kurie yra panaudojami skirtingose tinklo vietose. Šiame skyriuje yra atlikta saugumo sistemų ir metodų analitinė apžvalga ir detaliau aptarta jų paskirtis.

1.1.1. Incidentai informacinėse sistemose

Informacinių sistemų projektavimo pradžioje yra sunku nustatyti kokios saugumą užtikrinančios sistemos ir metodai turėtų užtikrinti pakankamą sistemos saugumą. Dėl to reikia nustatyti ir pasirinkti kelis komponentus ir metodus, kurie veiktų kartu ir užtikrintų apsaugą nuo grėsmių įvairovės. Kiekvienas apsaugos mechanizmas arba metodas gali užkirsti kelią tik tam tikram grėsmių sąrašui. Įvairių apsaugos mechanizmų taikymas kartu leidžia padidinti sistemos atsparumą nuo plataus grėsmių spektro (PTAC 2011a, 2011b). Toliau aptarsime populiariausius incidentus tinkle (CERT-LT ir CERT LitNET ataskaitų duomenimis) ir kokios yra apsaugos priemonės nuo jų.

1.1 lentelė. Populiariausi incidentai ir apsauga nuo jų
Table 1.1. Common incidents and defences

Incidentai	Aprašymas	Apsaugos būdas
Elektroninės paslaugos trikdymo ataka (angl. *DoS*)	Veiksmas, kuriuo siekiama sutrikdyti elektroninių ryšių tinklo ir (ar) informacinės sistemos darbą arba elektroninių ryšių tinklu teikiamas paslaugas.	Atakų atpažinimo ir prevencijos sistema, serverių PĮ atnaujinimas, ugniasienės panaudojimas, paketų filtravimas.
Kenksmingas programinis kodas (angl. *Malware*)	Programinė įranga ar jos dalis specialiai sukurta neteisėtai prisijungti ar sudaryti sąlygas neteisėtai prisijungti prie informacinės sistemos ar elektroninių ryšių tinklo, sutrikdyti ar pakeisti (taip pat pažeisti valdymą) informacinės sistemos ar elektroninių ryšių tinklo veikimą, sunaikinti, sugadinti, ištrinti ar pakeisti elektroninius duomenis, panaikinti ar apriboti galimybę naudotis elektroniniais duomenimis, sudaryti sąlygas neteisėtai pasisavinti ar kitaip panaudoti neviešus elektroninius duomenis tokios teisės neturintiems asmenims.	Antivirusinės programos naudojimas, antivirusinės programos duomenų bazės atnaujinimas, PĮ pataisymų naudojimas.
Nepageidaujamas elektroninis paštas (angl. *Spam*)	Elektroninio pašto laiškų tiesioginės rinkodaros tikslais siuntimas be elektroninio pašto naudotojo ir (ar) abonento išankstinio sutikimo, galintis turėti neigiamos įtakos elektroninio pašto informacinių sistemų funkcionalumui Lietuvos Respublikoje.	Atakų atpažinimo ir prevencijos sistema.
Neleidžiamasis prisijungimas (angl. *System Compromise/ Intrusion*)	Neteisėtas prisijungimas prie informacinės sistemos ar elektroninių ryšių tinklo.	Autentifikavimas ir slaptažodžių apsauga, operacinės sistemos apsauga, paketų filtravimas, ugniasienės.

1.1 lentelės pabaiga

Elektroninių duomenų klastojimas (angl. *Phishing*)	Sąmoningas elektroninių duomenų iškraipymas ar pakeitimas netikrais elektroniniais duomenimis	Autentifikavimas ir slaptažodžių apsauga.
Manipuliacija elektroniniais duomenimis (angl. *Spyware*)	Elektroninių duomenų pasisavinimas, platinimas, paskelbimas ar kitoks neteisėtas jų panaudojimas.	Autentifikavimas ir slaptažodžių apsauga.
Socialinė inžinerija (angl. *Social Engineering*)	Neteisėta prieiga prie informacinių išteklių, grindžiamų žmogaus psichologijos bruožų. Tai yra žmogaus pasitikėjimo išnaudojimas informacijos rinkimui ir sukčiavimui.	Edukacija, periodiniai mokymai.
Prievadų žvalga (angl. *Port Scaning*)	Piktavalis, siekiantis įsiskverbti į sistemą, ieško atidarytų tinklo prievadų.	Ugniasienės, leidžiančios nufiltruoti nepageidautiną srautą ir apsaugančios autorizuotą srautą įdiegimas, paketų filtravimo įdiegimas.

1.1.2. Apsaugos priemonės kompiuterių tinkle

Diegiant apsaugos priemonės privaloma laikytis taisyklių, kad viena iš apsaugos priemonių būtų pati veiksmingiausia. Sekančios apsaugos priemonės turi būti pridedamos prie esamų. Potencialus piktavalis turės sukompromituoti vieną apsaugos lygį tam, kad pasiekti sekantį. Apsaugos priemonių kiekio didinimas padaro sistemą labiau sudėtingą kompiuterių tinklo administratoriams. Dėl to saugumo priemonių kiekis turi balansuoti tarp kainos jų palaikymo sąnaudų ir naudos. Sistemos kaina neturėtų būti didesnė už apsaugos komponento nešamą naudą, kas deja yra sunkiai nustatoma (DID 2008; Lipiński 2014).

Apsaugos priemonės, grupuojamos pagal OSI ir TCP/IP tinklo modelį (Surman 2002; Holl 2003; IAIC 2012), gali būti šios:

Fizinis lygmuo:

- fizinė apsauga.

Tinklo lygmuo:

- paketų filtravimas;
- demilitarizuota zona (DMZ);
- ugniasienės;

- atakų atpažinimo ir prevencijos sistema.

Transporto lygmuo:

- virtualus privatus tinklas (VPN).

Programinės įrangos lygmuo:

- autentifikavimas ir slaptažodžių apsauga;
- operacinės sistemos apsauga;
- apsauga nuo kenksmingo kodo programų;
- tinklo auditas ir žiniaraščiai.

Fizinė apsauga

Fizinė apsauga yra skirta apsaugoti informacinės sistemos komponentus nuo vagystės, gaisro ar gamtos stichijos. Fizinės apsaugos pavyzdys būtų kompiuterio prirakinimas prie stalo, kompiuterių serverių patalpos rakinimas atskirame kambaryje ir pan. Atkreiptinas dėmesys, kad piktavalis gali prisiliesti prie informacinės sistemos fiziškai, vadinasi jis gali ją sukompromituoti.

Paketų filtravimas

Paketų filtravimo sistema blokuoja arba praleidžia paketų srautą remiantis prievado numeriu, IP adresu, protokolu arba kitokia informacija. Paketų filtravimas gali būti įgyvendintas panaudojus skirtingas apsaugos sistemas. Tai gali būti aparatiniai tinklo įrenginiai, tokie kaip maršrutizatoriai, komutatoriai, ugniasienės arba programinė įranga, kuri gali būti įdiegta į pačią informacinę sistemą.

Ugniasienės

Ugniasienė skirta apsaugoti kompiuteryje esančius duomenis nuo įsilaužėlių. Ugniasienės būna aparatinės ir programinės. Jos gali dirbti dviem režimais: filtruojant srautą – leidžiama viskas pagal nutylėjimą, bet blokuojami tam tikri srautai ir blokuojant srautą – kai pagal nutylėjimą viskas yra blokuojama, o praleidžiami tik tam tikri srautai.

Demilitarizuota zona (DMZ)

Tinklo dalis, kuri yra už vidinės tinklo ribos, bet yra pajungta į ugniasienę, vadinama demilitarizuota zona. Demilitarizuotoje zonoje gali būti viešai prieinami serveriai. Dažnai tai yra vadinama servisų potinkliu arba perimetro tinklu.

Atakų atpažinimo sistema

Idealiu atveju ugniasienės ir *proxy* serveriai blokuoja atakas arba kenkėjišką kodą nuo patekimo į vidinį tinklą, bet atakų atpažinimo sistemos panaudojimas leidžia atpažinti dar iki šiol nežinomas atakas, kurios gali patekti į tinklą. Atakų atpažinimo sistema yra paremta atakų požymių arba anomalijų tinkle atpažinimu bei administratoriaus informavimu apie tinkle atsirandančius įvykius.

Virtualus privatus tinklas (VPN)

Bendrovės, kurios dalinasi bylomis arba apsikeičia konfidencialia informacija tarp padalinių, dažniausiai naudoja dedikuotus tinklo sujungimus siūlomus telekomunikacijų bendrovių. Tokie dedikuoti tinklai atrodo kaip taškas-taškas sujungimas tarp bendrovės padalinių ir yra pakankamai apsaugoti, bet toks sprendimas yra labai brangus. Labai dažnai yra naudojami virtualūs privatūs tinklai leidžiantys realizuoti pigų ir saugų ryšį per viešąjį tinklą. Virtualus privatus tinklas – tai tinklas, kuris naudoja viešojo interneto infrastruktūrą bei leidžia saugiai prijungti unikalius vartotojus prie bendrovės tinklo. Virtualūs privatūs tinklai naudoja autentifikavimą, autorizuoja vartotojus ir šifruoja siunčiamą tinklo srautą.

Autentifikavimas ir slaptažodžių apsauga

Po to kai informacinė sistema yra fiziškai apsaugota reikia ją apsaugoti iš vidaus. Viena paprasta, bet efektyvi priemonė yra slaptažodžio apsaugos politika, kuri reikalauja iš vartotojo pasirinkti sudėtingus ir sunkiai atspėjamus slaptažodžius, laikyti juos paslaptyje bei reguliariai atnaujinti. Būtina naudoti skirtingus slaptažodžius pradedant nuo darbastalio rakinimo iki individualios programinės įrangos ir duomenų bazių. Tačiau pasinaudojus modernia technika, slaptažodžio apsauga gali būti nulaužta, o slaptažodžiai atskleisti. Autentifikavimas – tai procesas, kurio funkcija yra tapatybės nustatymas, t. y. nustatyti, kad vartotojas, programinė įranga arba kompiuteris yra iš tikrųjų tas kuo jis save pristato. Bazinis autentifikavimas yra paremtas tam tikros informacijos patikrinimu, kurią žino vartotojas, pvz. vartotojas/slaptažodis. Galimas ir biometrinės-fizinės technologijos panaudojimas, kuris vartotoją leidžia identifikuoti pagal fiziologines ar elgsenos charakteristikas.

Operacinės sistemos apsauga

Kitas informacinės sistemos apsaugos iš vidaus būdas yra nuolatinis operacinės sistemos atnaujinimų diegimas. Dažniausiai atnaujinimai yra išleidžiami saugumo spragoms ištaisyti. Atnaujinimo procesas yra vartotojo arba administratoriaus atsakomybė. Taip pat operacinės sistemos nenaudojamų servisų atjungimas leidžia padidinti sistemos saugumą ir atsparumą atakoms.

Apsauga nuo kenksmingo kodo programų

Dažniausias ginklas prieš kompiuterinius virusus yra antivirusinės programos, kurios atlieka sistemos skanavimą ieškant virusų duomenų bylose arba elektroniniuose laiškuose. Dažniausiai virusai yra *.exe (vykdomasis kodas) arba *.zip (archyvuotos bylos) bylose. Ieškant kenkėjiško programinio kodo antivirusinė sistema naudoja kelis metodus, pvz. bylų kontrolinės sumos patikrinimas ir palyginimas, viruso pirštų antspaudų paieška ir kt. Antivirusinė sistema atpažįsta virusą, identifikuoja jo vietą, vėliau jį ištrina arba patalpina į specialiai išskirtą atminties vietą, kur kenkėjiškas kodas negali būti įvykdytas. Pažymėtina, kad ugniasienės ir atakų atpažinimo sistemos nėra skirtos virusų skanavimui ir kovai su jais. Tuo tarpu daugelis šiuolaikinių ugniasienių turi antivirusinės sistemos funkcijas, kas dar labiau padeda kontroliuoti visus tinklo duomenų srautus. Norint apsaugoti kompiuterių tinklą nuo žalingos virusų veiklos būtina įdiegti antivirusinę sistemą.

Tinklo auditas ir žiniaraščiai

Tinklo auditas – tai procesas fiksuojantis kuris kompiuteris prieina prie tinklo ir kokius resursus naudoja. Informacija yra surašoma į elektroninius žiniaraščius. Dažniausiai administratoriai analizuoja ugniasienių ir atakų atpažinimo sistemų žiniaraščius. Turint ilgalaikę įrašų istoriją galima aptikti įtartiną veiklą tinkle. Tokia informacija leidžia pamatyti kas bandė įsilaužti į tinklą ir imtis atitinkamų veiksmų. Žiniaraščių analizė, tai viena iš neatsiejamų saugumą užtikrinančių priemonių.

1.2. Tinklo saugumo modeliavimo metodų analitinė apžvalga ir metodo pasirinkimas

Spartus technologijų vystymasis įtakoja kompiuterių tinklo bei informacinių sistemų projektavimo procesą. Tokios sistemos yra kuriamos keliant didelius reikalavimus sistemos spartai bei patikimumui. Suprojektuota ir įgyvendinta sistema turi atitikti jai keliamus saugumo reikalavimus. Saugumo modeliavimas yra svarbus tokių sistemų projektavimo etapas (Heidari 2006). Jis gali gerokai sumažinti informacinės sistemos projektavimo kaštus ir trukmę, kadangi leidžia analizuoti sistemos elgseną pradedant nuo projektavimo pradžios iki sistemos galutinio įgyvendinimo (Paulauskas *et al.* 2009).

Daugelyje užsienio ir Lietuvos autorių darbų yra analizuojamas sistemos saugumas, kai jį paveikia vienas arba keli saugumo incidentai. (Ramanauskaitė *et al.* 2011a, 2012) darbuose yra analizuojama kaip atsisakymo aptarnauti ataka įtakoja kompiuterių sistemos komponentus, tokius kaip tinklo pralaidumas,

kompiuterio sistemos atmintis ir centrinis procesorius. Sistemos modeliavimui yra naudojamas matematinis ir programinis modeliai. (Ramanauskaitė *et al.* 2009, 2010, 2011b) darbuose apžvelgiami egzistuojantys TCP SYN, SYN ir DoS atakų tipai bei siūlomas stochastinis modelis atakoms modeliuoti. Jis leidžia atsižvelgi į teisėtą sistemos srautą, galimą atakos galingumą, aukos naudojamų apsaugų savybes ir gali būti naudojamas ne tik TCP SYN, bet ir kitoms atminties išnaudojimo DoS atakoms modeliuoti. (Goranin *et al.* 2008, 2009) straipsniuose siūlomas genetiniu algoritmu pagrįstas modelis, skirtas žinomų ir perspektyvių interneto kirminų plitimo greičiams nustatyti po prisisotinimo fazės. Algoritmas pagrįstas žinomų kirminų plitimo strategijų ir susijusių plitimo greičių analize ir pateikiamas kaip sprendimų medis. Modeliai paremti genetiniais algoritmais. (Kajackas *et al.* 2011a, 2011b) darbuose parengtas modelis, kuriuo galima įvertinti kibernetinių atakų poveikį Lietuvos interneto tinklo infrastruktūrai ir nustatyti labiausiai pažeidžiamus mazgus ir linijas tinkle. Sudarytas virtualus Lietuvos interneto tinklo modelis bei paruošti ir išbandyti skirtingi kibernetinių atakų scenarijai. Tinklo modelis yra sukurtas OPNET programinėje įrangoje ir panaudojant matematinius algoritmus. (Puniškis *et al.* 2005, 2006, 2007; Laurutis 2003) savo darbuose pateikia informacinių sistemų modelius paremtus neuroniniais tinklais, kur modeliuoja kenksmingo programinio kodo aptikimą ir nepageidaujamo elektroninio laiško filtravimą.

Informacinės sistemos saugumas apibūdinamas keliais parametrais, pagrindiniai iš jų yra: konfidencialumas, vientisumas ir pasiekiamumas. Šiuo metu žinoma eilė informacinių sistemų saugumo modeliavimo metodų, kurie taikomi projektuojant naujas sistemas bei vertinant jau eksploatuojamų sistemų saugumą, tai būtų tikimybiniai, Markovo procesų, Petri tinklų ir kiti metodai (Nicol *et al.* 2004). Šio poskyrio tikslas yra išnagrinėti žinomus modeliavimo metodus, aptarti jų galimybes, privalumus bei trūkumus ir pasirinkti metodą, tinkamiausią pagrindinių saugumo veiksnių įvertinimui.

1.2.1. Tikimybiniai metodai

Tikimybių teorija gali būti taikoma patikimumo ir pasiekiamumo modeliavimui priėmus tam tikras prielaidas, kad sistemos komponentų įvykių tikimybės yra nepriklausomos (Kbar 2009; Haimes 2005; Elahi *et al.* 2011).

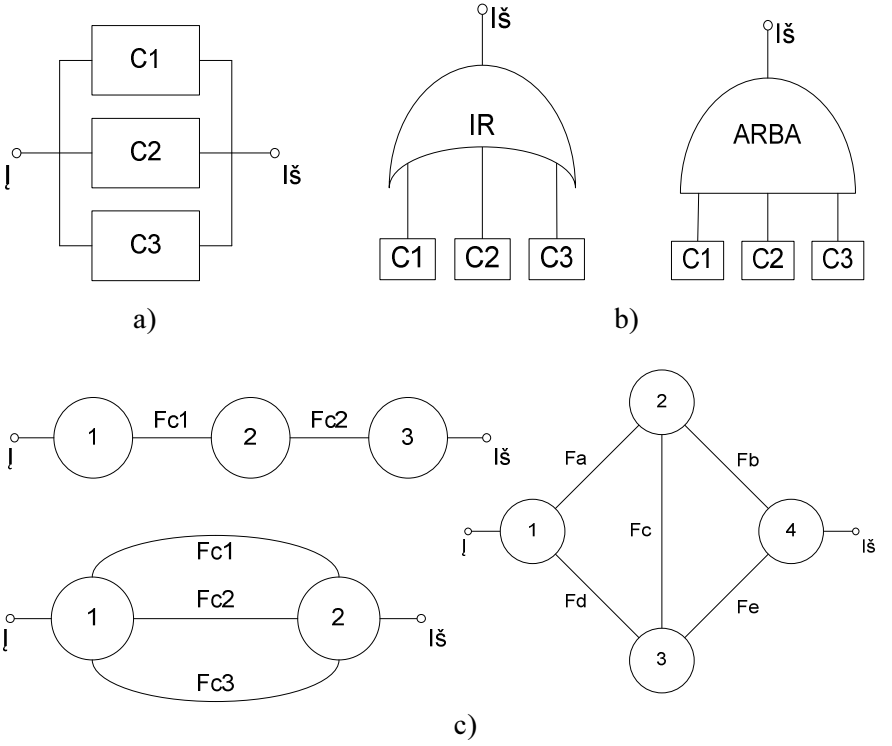
Universali charakteristika, tinkanti aprašyti kaip diskrečiuosius, taip ir tolydžiuosius dydžius, yra pasiskirstymo funkcija. Pasiskirstymo funkcija arba integraliniu tikimybių skirstiniu vadinama funkcija *F(X),* atitinkanti tikimybę *P*, kad atsitiktinis dydis *X* įgis vertes mažesnes už *x* intervale nuo 0 iki *x*:

$$F(X) = P(X < x). \tag{1.1}$$

Kai kuriais atvejais atsitiktinis dydis aprašomas nesinaudojant pasiskirstymo funkcija ar tikimybės tankiu, o imant tam tikras jo skaitines charakteristikas. Paprasčiausios ir svarbiausios atsitiktinio dydžio skaitinės charakteristikos yra matematinė viltis ir dispersija.

Sistemos pasiekiamumas yra apskaičiuojamas analogiškai, tik vietoje negendamumo tikimybės *P(t)* yra naudojama pasiekiamumo tikimybė *A(t)*. Pasiekiamumas – tai tikimybė, kad sistema atliks savo funkcijas užduotame laiko intervale.

Tinklo sistemos paprastai susideda iš kelių komponentų, tai gali būti maršrutų parinktuvai, komutatoriai, serveriai ir t. t. Komponentus sudaro smulkesni komponentai, pvz. atmintis, procesoriai, valdikliai, kaupikliai ir t. t. Taigi visos sistemos negendamumas priklauso nuo visų komponentų negendamumo tikimybių. Sistema gali sugesti, kai vienas, keli arba visi komponentai sugenda.



1.1 pav. Sistemos komponentų patikimumo grafinis atvaizdavimas: a) patikimumo diagramos; b) patikimumo medis; c) patikimumo grafai

Fig. 1.1. System component realiability graphs: a) reliability diagram; b) reliability tree; c) reliability graph

Modeliuojant tokią sistemą tikimybiniais metodais priimama prielaida, kad komponentų gedimų tikimybės nepriklausomos ir neįtakoja kitų komponentų gedimo tikimybių. Informacinės sistemos gali būti atvaizduojamos grafiškai (1.1 pav.). Yra keli populiarūs grafinio atvaizdavimo metodai, pavyzdžiui gedimų medis, patikimumo diagramos, patikimumo grafai. Šie grafinio atvaizdavimo metodai labai mažai skiriasi tarpusavyje, yra tiktai daugiau ar mažiau populiarūs.

Grafiškai pavaizduota sistema leidžia įvertinti komponentų gedimų tikimybes ir jų išsidėstymą sistemoje. Sistema yra veikianti kol yra bent vienas kelias tarp įėjimo ir išėjimo.

Tikimybių teorijos metodai yra tinkami naudoti statinių sistemų modeliavimui, bet netinka dinaminėms sistemoms. Realybėje yra sutinkamos dinaminės sistemos. Taip pat tikimybių teorijos prielaidos dažniausiai netinka realioms sistemoms, kadangi sistemos komponenčių gedimų tikimybės gali būti priklausomos, t. y. vienas įvykis gali įtakoti kito įvykio tikimybę.

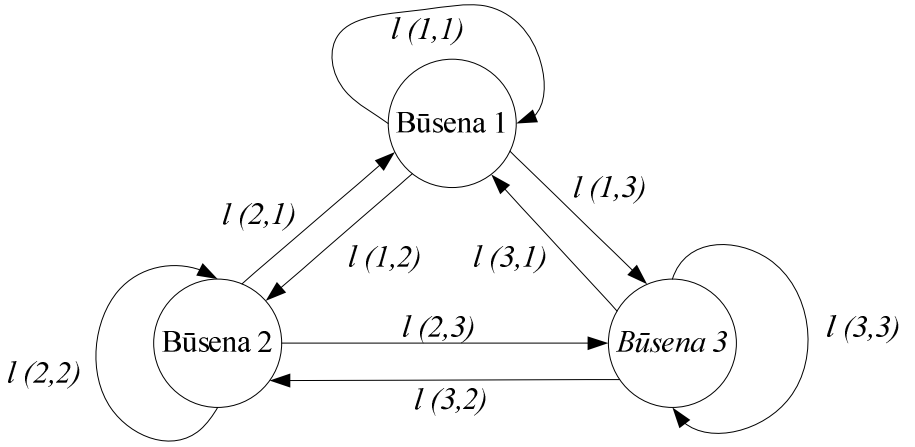
1.2.2. Markovo procesai

Dinaminių sistemų modeliavimui yra taikomi Markovo procesai (Pukite *et al.* 1998; Bakhoum 2011; Alsubhi *et al.* 2012; Haverkort *et al.* 2001; Bolch *et al.* 1998; Trivedi 2001). Sistema yra padalinama į būsenas, perėjimas tarp būsenų priklauso tik nuo esamos būsenos ir nepriklauso nuo praeities įvykių. Tam, kad galima būtų taikyti Markovo procesus turi būti priimtos tam tikros prielaidos, t. y. perėjimas tarp sistemos būsenų tūri būti pasiskirstęs pagal eksponentinį dėsnį. Eksponentinio pasiskirstymo funkcija yra be atminties, tai reiškia, kad vidutinis įvykių atsiradimo dažnis laiko intervale yra pastovus ir nepriklauso nuo jo padėties laiko ašyje, o priklauso tiktai nuo intervalo ilgio.

Norint aprašyti Markovo procesą reikia:

- nurodyti visas būsenas, kuriose gali būti sistema;
- sudaryti sistemos būsenų grafą ir jame nurodyti visus galimus perėjimus iš vienos būsenos į kitą;
- kiekvienam perėjimui nurodyti atitinkamą įvykių srauto intensyvumą $\lambda_{ij}(t)$, kuris perveda sistemą iš vienos būsenos S_i į kitą būseną S_j;
- nurodyti pradinę sistemos būseną, kai $t = 0$.

Tuo atveju, kai sistemoje vykstantis nepertraukiamas procesas yra Markovo, jį galima aprašyti paprastomis diferencialinėmis lygtimis, kuriose nežinomomis funkcijomis yra būsenų tikimybės. Sistemos būsenų grafas, esant 3 būsenoms, gali būti atvaizduotas grafiškai (1.2 pav.).

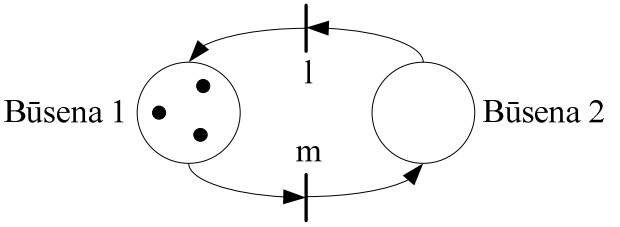


1.2 pav. Sistemos, sudarytos iš trijų būsenų, grafas
Fig. 1.2. Graph of system with three states

Markovo grandines galima taikyti sistemos patikimumo, pasiekiamumo, našumo ir išliekamumo modeliavimui. Vienas iš Markovo grandinių trūkumų yra tas, kad sudarant grandinę turi būti apibrėžtos visos galimos sistemos būsenos. Realių sistemų būsenų aibė gali būti labai didelė, dėl to Markovo grandinė tampa labai sudėtinga. Kuomet sistema turi daug (šimtus ar tūkstančius) būsenų be kompiuterio sudaryti būsenų grafą ir atitinkamą algebrinių lygčių sistemą praktiškai neįmanoma. Jei lygčių sistemą ir turėtume, tai gauti analitinį sprendinį taip pat retai pavyksta. Antras trūkumas yra modelio nelankstumas, dėl to realios sistemos elgsenos atkūrimas dažniausiai yra komplikuotas.

1.2.3. Petri tinklai

Alternatyva Markovo grandinėms yra Petri tinklai. Petri tinklai buvo sukurti dinaminių sistemų modeliavimui. Jais galima modeliuoti tiek programinės įrangos, tiek tinklų sistemų patikimumą (Sallhammar *et al.* 2005; Yang *et al.* 2010; Nianhua *et al.* 2011; Marsan *et al.* 1995; Hirel *et al.* 2000). Petri tinklų panaudojimas leidžia išvengti minėtų Markovo grandinių trūkumų. Šiuo atveju modelio dydis neišauga dėl komponentų kiekio, kadangi yra naudojamos ne globalios, o lokalios sistemos būsenos. Be to, Markovo grandinėse perėjimas tarp būsenų turi būti pasiskirstęs pagal eksponentinį dėsnį, kas ne visada tinka realioms sistemoms.



1.3 pav. Sistemos, sudarytos iš trijų komponentų, Petri tinklo modelio pavyzdys
Fig. 1.3. Petri network model example of system with three state

Petri tinklai yra sudaromi pasinaudojus tokiais elementais, kaip pozicijos, veiklos, žetonai ir lankai. Pozicijos atvaizduoja sistemos būseną. Veiklos atvaizduoja įvykių prigimtį, įvykiai gali būti momentiniai, įvykiai su vėlinimu ir įvykiai pasiskirstę laike pagal pasirinktą pasiskirstymo dėsnį. Žetonai modelyje atvaizduoja sąlyginius objektus, kurie realioje sistemoje gali būti sistemos komponentais. Lankai – tai keliai, kuriais žetonai gali judėti modelyje. 1.3 paveiksle parodytas sistemos, sudarytos iš trijų komponentų, Petri tinklas. Kiekvienas iš komponentų gali sugesti su ta pačia tikimybe λ, komponentas atstatomas su μ tikimybe. Petri tinkle sistemos būsena yra nuskaitoma iš pozicijų, kurios yra aprašomos lokaliai, dėl to nuo komponentų skaičiaus modelis nedidėja (Nianhua *et al.* 2011).

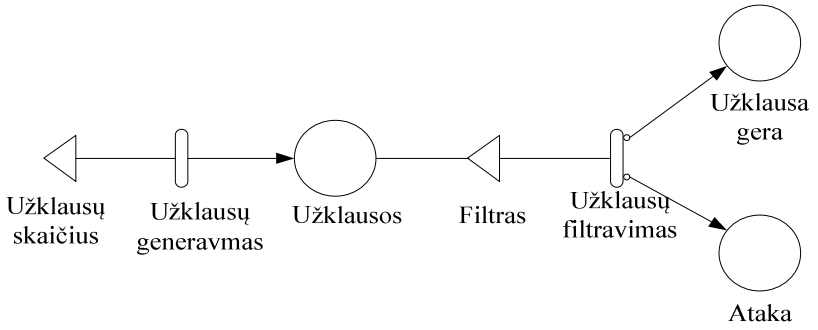
Petri tinklus yra lengviau sudaryti, modifikuoti ir analizuoti negu Markovo grandines. Petri tinklai – tai patogi grafinė, aukšto lygio kalba aprašanti sistemos elgseną. Yra keli Petri tinklų tipai (su įvairiais funkcijų plėtiniais), kurie taikomi sprendžiant specifinius uždavinius. Laiką įvertinantis Petri tinklas (angl. *Petri Net with Time*) – laikas buvo įvestas norint modeliuoti sąveikas tarp kelių procesų atsižvelgiant į jų pradžią ir pabaigą. Stochastiniai Petri tinklai SPN (angl. *Stochastic Petri Nets*) – tai laiką įvertinantys tinklai, kurių būsenos keičiasi nepriklausomai, o perėjimų vėlinimai yra atsitiktiniai ir pasiskirstę pagal eksponentinį dėsnį. Apibendrintas stochastinis Petri tinklas GSPN (angl. *Generalized Stochastic Petri Net*) įvertina dviejų rūšių sistemos būsenų kaitą, nepriklausančią nuo laiko ir priklausančią nuo laiko (pasiskirsčiusią eksponentiniu dėsniu). Neuždelsti perėjimai turi pirmenybę prieš priklausančius nuo laiko perėjimus. Be to neuždelsti perėjimai gali turėti prioritetus vienas kito atžvilgiu. Egzistuoja ir daugiau Petri tinklų rūšių. Šio formalaus metodo paplitimas parodo jo paprastumą ir universalumą.

1.2.4. Stochastinės veiklos tinklai

Kompiuterių ir tinklo sistemų patikimumo, pasiekiamumo, našumo ir išliekamumo įvertinimo modeliavimo metodai vystėsi tiek pat ilgai, kiek ir

pačios sistemos. Vieni metodai, kurie turėjo trūkumų arba apribojimų, buvo pakeičiami kitais metodais. Vienas iš labiausiai patobulintų ir lanksčių modeliavimo metodų yra stochastinės veiklos tinklai, tai yra patobulintas Petri tinklų metodas. Stochastiniai veiklos tinklai (angl. *Stochastic Activity Network – SAN*) yra lanksti ir lengvai taikoma stochastinių Petri tinklų atmaina (Beaudet *et al.* 2006; Pinsky *et al.* 2001; Meyer *et al.* 1985; Deavours *et al.* 2002).

Stochastinės veiklos tinkluose yra galimi papildomi tinklo modeliavimo objektai – įėjimo ir išėjimo vartai, leidžiantys kontroliuoti įvykių pradžią ir pabaigą priklausomai nuo aprašytų sąlygų. Stochastiniai veiklos tinklai yra stochastinių veiklos tinklų AN (angl. *Activity Network*) išplėtimas, panašiai kaip stochastiniai Petri tinklai yra klasikinių Petri tinklų išplėtimas. Stochastinės veiklos tinkluose yra galimas matematinis sąlygų aprašymas panaudojant tokius veiksmus, kaip sudėtis, atimtis, palyginimas daugiau, mažiau, prilyginimas nuliui. Pasinaudojant stochastinės veiklos tinklais galima sukurti sudėtingus ir kompleksinius įvairių sistemų modelius. Veikla gali turėti kelis jos baigties atvejus. Įėjimo vartai turi įjungimo sąlygą ir funkciją, kuri kontroliuoja veiklų vykdymą. Išėjimo vartai turi tik funkciją, kuri nusako kaip keičiasi žyma įvykdžius veiklą. Vartų naudojimas suteikia daugiau lankstumo, todėl SAN yra funkcionalesni už Petri tinklus (Garšva *et al.* 2011; Garšva *et al.* 2006).



1.4 pav. Stochastinės veiklos tinklų pavyzdys
Fig. 1.4. Example of stochastic activity network model

Stochastinės veiklos tinklo pavyzdys yra pateiktas 1.4 paveiksle. Tai yra tinklo sistemos vartotojo modelis, kuris realizuotas kaip vartotojo užklausų generatorius. Modelyje yra apibrėžtos tokios modelio sąlygos, kaip sugeneruotų užklausų kiekis, užklausų pasiskirstymo dėsnis, užklausų filtravimas pagal tam tikrus kriterijus modeliuojant ugniasienę, užklausų pasiskirstymas pagal prigimtį – ataka ar užklausa.

Pasinaudojant 1.4 paveiksle pateiktu modeliu galima įvertinti pagrindines saugumo charakteristikas: konfidencialumą, vientisumą bei pasiekiamumą.

Kiekviena užklausa gali vienaip ar kitaip įtakoti sistemą, tai gali būti visos sistemos arba jos dalies gedimas, sistemos duomenų atskleidimas, įtakojantis konfidencialumą arba kito pobūdžio ataka. Patikimumas $P_{(s=1)}$ – tai tikimybė, kad informacinė sistema bus normaliame būvyje ($s = 1$). Vientisumas – tai tikimybė, kad visą sistemos eksploatavimo laiką T sistema nepereis iš normalaus pradinio būvio ($s = 1$) į kokį nors kitą. Modelis yra paremtas įvairiausių įvykių tikimybėmis, kurios yra pasiskirsčiusios laike.

Iš atliktos modeliavimo metodų apžvalgos galime padaryti išvadas. Tikimybių teorijos metodai tinka tik statinių sistemų charakteristikų modeliavimui. Markovo procesai tinkami modeliuoti dinamines sistemas. Tačiau panaudojant Markovo procesus sistemos modelis greitai išauga dėl komponenčių skaičiaus padidėjimo, nes yra modeliuojamos globalios būsenos. Petri tinklai yra gera alternatyva Markovo grandinėms. Privalumas, kad modelis neišauga dėl komponenčių skaičiaus padidėjimo, kadangi vietoje globalių yra modeliuojamos lokalios būsenos. Petri tinklus yra lengviau sudaryti, modifikuoti ir analizuoti negu Markovo grandines. Petri tinklai yra apriboti tokiomis sąlygomis kaip +, -, > ir nulio palyginimas. Yra sunku modeliuoti sudėtingas sistemas. Stochastinės veiklos tinklas, tai Petri tinklas su išplėstomis galimybėmis. Yra galimybė aprašyti įvykių pradžios, pabaigos sąlygas, pasirinkti laikines ir momentines veiklas, jų pasiskirstymo dėsnį. Būsenų keitimo galimybė yra priklausoma nuo įvykių. Galimi papildomi tinklo aprašymo objektai: perėjimo sąlygos, įėjimo ir išėjimo vartai. Apžvalga parodė, kad iš aptartų modeliavimo metodų, dinaminių sistemų saugumo charakteristikų modeliavimui labiausiai tinka stochastinės veiklos tinklų metodas. Stochastinės veiklos tinklai leidžia įvertinti skirtingus įvykių pasiskirstymo dėsnius. Stochastinių veiklos tinklų modeliai yra funkcionalesni už Petri tinklų modelius, be to jie gali būti lengvai tobulinami ir redaguojami.

1.3. Tinklo srautų statistikos panaudojimas

Tinklo srautų analizė pateikia informaciją, reikalingą tinklo srautų modeliavimui, srauto valdymui, tinklo apkrovų planavimui, prognozei, anomalijų ir atakų atpažinimui.

Tinklo srauto modeliavimas yra paremtas informacijos apie tinklo srautus rinkimu (Barakat *et al.* 2002, 2003). Straipsnio autoriai sudarė tinklo srauto modelį neapkrautiems stuburinio tinklo sujungimams. Modelis yra paremtas Puasono procesu srauto lygmenyje. Jis leidžia surasti tinklo būsenos pasikeitimus trumpame laiko intervale tik panaudojus ribinių tinklo prievadų statistinę analizę. Modelis padeda tinklo srauto maršrutizavimo metu įvertinti prievadų apkrovimą ir rasti optimaliausią maršrutą.

A. Lakhina ir kt. (Lakhina *et al.* 2003, 2004) parodo pilno srauto rinkinio (siuntėjas ir gavėjas) analizę stuburinio akademinio tinklo prievaduose. Autoriai naudoja principinių komponentų analizę. Jie rado, kad tinklas su daugiau kaip šimtais srautų gali būti modeliuojamas panaudojant mažą kiekį nepriklausomų komponentų arba matmenų. Principinė komponentų analizė leidžia dekomponuoti srauto struktūrą į tris pagrindines dalis: dažną periodinį aktyvumą, trumpus pikus ir pastovų triukšmą. Modelis leidžia spręsti ir analizuoti tokias problemas: įvertinti srauto matricą, atpažinti anomalijas, prognozuoti srautą.

Staigus ir tikslus anomalijų atpažinimas yra vienas iš svarbesnių didelio kompiuterių tinklo uždavinių. (Barford *et al*. 2001, 2002) straipsnių autoriai bando atpažinti anomalijas panaudodami srauto formos filtrus, paremtus SNMP ir *NetFlow* protokolų surinkta informacija. Jų gauti rezultatai parodo, kad šis metodas yra efektyvus detektuojant įprastus ir anomalijų srautus. Srauto formos filtrų panaudojimas taip pat aprašomas (Crovella *et al*. 2003; Abry *et al.* 1998). Straipsnyje (Yegneswaran *et al*. 2003) autoriai tyrinėja tinklo aktyvumą remdamiesi ugniasienės įrašais. Tinklo atakos aktyvumas gali būti pastebėtas pamačius didelį prievadų skanavimo dažnį tinkle. Tinklo srautų informacija taip pat gali būti panaudota prievadų skanavimo aptikimui. Straipsnis parodo saugumo incidentų aptikimo galimybes panaudojant *NetFlow* protokolą.

Tinklo srautų statistika taip pat gali būti panaudota tinklo sujungimų apkrovų ir perpildymų nustatymui ir jų prognozei ateityje. (Papagiannaki *et al.* 2003, Aussem *et al.* 1998) straipsnio autoriai siūlo tinklo sujungimų prognozės metodologijas. Pirmame straipsnyje parodoma, kad panaudojant SNMP surinktą statistiką gautos stuburinio tinklo srauto charakteristikos atvaizduoja matomas ilgo laikotarpio tendencijas, jose matomi ryškūs pasikartojimai. Pasiūlyta metodologija leidžia numatyti tinklo susijungimų ateities apkrovą.

Statistinė paketų analizės problema didelio našumo tinkluose yra aprašyta straipsnyje (Zhang *et al.* 2003, Li *et al.* 2003) *NetFlow*, SNMP arba Linux srautų įrankiai yra naudojami didelės greitaveikos prievaduose ir didelio duomenų kiekio surinkimui. Surinkta agreguota tinklo srautų statistika leidžia sumažinti didelės saugojimo atminties poreikį, o surinkta informacija yra pakankamai informatyvi tinklo srautų analizei.

Surinkta tinklo statistinė informacija gali būti panaudota tinklo srauto generavimo modeliuose (Vishwanath *et al.* 2006; Sommers *et al.* 2004; Cao *et al.* 2002, 2004a, 2004b; Cleveland *et al.* 2000).

1.4. Pirmojo skyriaus išvados ir disertacijos uždavinių formulavimas

1. Kompiuterių tinklų dinaminių saugumo charakteristikų modeliavimui pasirinktas stochastinės veiklos tinklų metodas, kuris leidžia pasirinkti papildomus tinklo aprašymo objektus bei įvertinti skirtingus įvykių pasiskirstymo dėsnius. Kuriamas modelis gali būti lengvai tobulinamas ir redaguojamas, o jo dydis nuo to nesikeičia.

2. Apžvelgtoje literatūroje nėra pateikiama srautų ir paketų pasiskirstymo tinkluose statistinių duomenų surinkimo ir analizės metodika. Nėra informacijos apie atskirų incidentų įtaką konfidencialumui, vientisumui ir pasiekiamumui.

3. Modeliuojant tinklo duomenų srautą dažniausiai naudojamas paprasčiausias srautas su eksponentiniu pasiskirstymu. Priimama, kad paketų atėjimo laikai pasiskirstę pagal Puasono dėsnį.

4. Iki šiol nėra sukurta informacinės sistemos išliekamumo modeliavimo metodika, kuri leistų pritaikyti sistemos sukompromitavimo tikimybes remiantis grėsmės tipu, sunkumu ir saugumo mechanizmų rinkiniu. Taip pat iš apžvelgtų darbų matyti, kad nėra tokio modelio, kuris įvertintų bendrą sistemos saugumą turint informaciją apie bendrą įvairių incidentų tipų statistiką.

Darbo tikslui pasiekti, darbe reikia spręsti šiuos uždavinius:

1. Išanalizuoti tinkle egzistuojančius incidentus ir nustatyti jų įtaką saugumo sudedamosioms dalims – konfidencialumui, vientisumui ir pasiekiamumui.
2. Išanalizuoti informacinių sistemų saugumo modeliavimo metodus.
3. Atlikti kompiuterių tinklo duomenų srautų statistinę analizę, rasti TCP ir UDP paketų pasiskirstymo dėsnius.
4. Informacinių sistemų išliekamumo įvertinimo modelio sudarymas.
5. Gauti informacinės sistemos saugumo lygio vertinimo charakteristikas.

2

Kompiuterių tinklo srautų statistinė analizė

Šiame skyriuje pateikiami akademinio kompiuterių tinklo srautų statistinės analizės rezultatai, gauti panaudojus tinklo srauto rinkimo protokolą *NetFlow*. Statistiniai rezultatai vaizduojami grafine forma, kuri atskleidžia kompiuterių tinklo srautų pasiskirstymo tendencijas.

Kompiuterių tinklo duomenų srautų modeliavimui reikia turėti informaciją apie kompiuterių tinklo charakteristikas tokias kaip duomenų paketų pasiskirstymo dėsnius. Modeliavimui būtina turėti duomenų paketų pasiskirstymo dėsnio matematinę išraišką (Paulauskas *et al.* 2009), bet pagrindine problema lieka atsakymas į klausimą ar statistinės formulės atvaizduoja realią situaciją. Žemiau pateiktas tyrimas remiasi Vilniaus Gedimino technikos universiteto (toliau – VGTU) kompiuterių tinkle surinkta informacija. Pagrindinis tinklo srautų informacijos šaltinis yra stebėjimo sistema, kuri naudoja *NetFlow* protokolą. *NetFlow* protokolas yra skirtas informacijos apie srautus rinkimui tinkluose su didele apkrova. Informacijos rinkimas naudojant *NetFlow* protokolą yra ribotas, todėl šiame tyrime teks iš jos išrinkti kuo daugiau su tinklo atakomis susijusių duomenų, kuriuos panaudosime sudarydami atakų atpažinimui skirtus tinklo modelius.

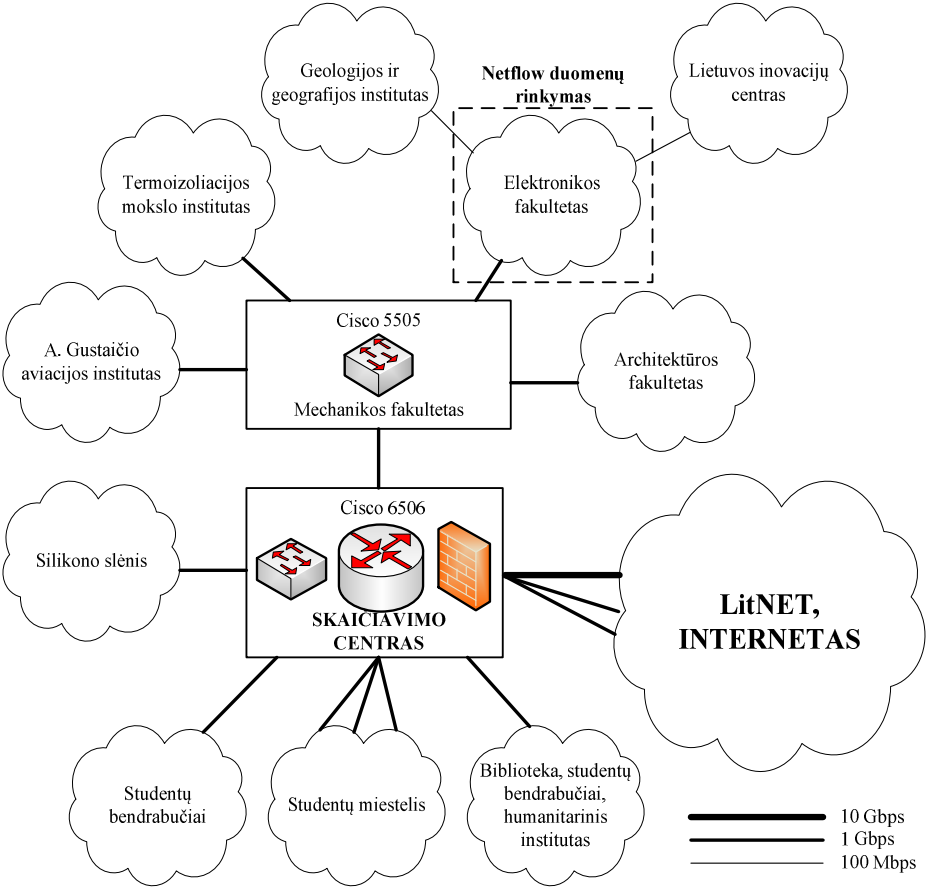
Šiame skyriuje išanalizuotas kompiuterių tinklo srautas Elektronikos fakultete ir atlikta srauto duomenų srautų statistinė analizė. Tinklo srautai buvo

išanalizuoti ir sugrupuoti pagal protokolus, pasirodymo laiką, paketų dydį bei unikalų IP adresą, o vėliau panaudoti tinklo analizei ir modelio sudarymui.

Gautos informacijos apimtys ir jos apdorojimo būdų kiekis yra labai dideli, dėl to yra pateikiama tik TCP protokolo analizė. Šio skyriaus medžiaga publikuota autoriaus straipsniuose (Garšva, Paulauskas, Gražulevičius, Gulbinovič 2012, 2013, 2014).

2.1. Akademinis kompiuterių tinklas

VGTU kompiuterių tinklas sujungia universiteto fakultetus ir skyrius bei jungiasi su globaliu pasaulio tinklu (2.1 pav.).



2.1 pav. VGTU kompiuterių tinklas
Fig. 2.1. Scheme of VGTU computer network

Universiteto kompiuterių tinklas yra sukurtas naujausių transmisijos ir valdymo technologijų pagrindu, panaudojant optinę prieigą. Universiteto kompiuterių tinklas yra Lietuvos akademinio tinklo ir LitNet tinklo dalis, kuri apjungia apie 3 250 kompiuterių.

Visi VGTU tinklo duomenų srautai yra kontroliuojami vienu Cisco 6506 maršrutizatoriumi, per kurį iš LitNet tinklo internetas ateina į visus universiteto tinklo mazgus. Per pagrindinį universiteto maršrutizatorių metų bėgyje perduodama apie 700 TB duomenų informacijos abiem kryptimis. VGTU LitNet tinklo mazgai turi 32 B klasės IP adresų potinklius, tinkle yra 8 160 viešųjų IP adresų. Universiteto tinklas yra padalintas į 29 potinklius, skirtus atskiriems padaliniams ir fakultetams. Elektronikos fakultetas prižiūri tam tikrą universiteto kompiuterių tinklo dalį, kuri yra pažymėta schemoje (2.1 pav.).

IP srauto informacijos kaupimui naudojamas *NetFlow* protokolas sukurtas Cisco Systems Incorporated. Šiandien *NetFlow* tapo pramonės standartu duomenų srauto stebėjimui.

NetFlow protokolas eksportuoja tinklo srautų duomenis iš maršrutizavimo įrenginių (*NetFlow* 2012). Tinklo srautų stebėjimo ir analizės sistemos, paremtos šia technologija, sudedamosios pagrindinės dalys yra:

- *NetFlow* šaltinis. Tai gali būti vienas įrenginys, kuris valdo visos organizacijos srautus. Visa informacija apie tinklo srautus yra eksportuojama iš pagrindinio VGTU kompiuterių tinklo maršrutizatoriaus Cisco 6506 serijos;
- *NetFlow* rinktuvas;
- analizatorius, skirtas surinktų duomenų apdorojimui.

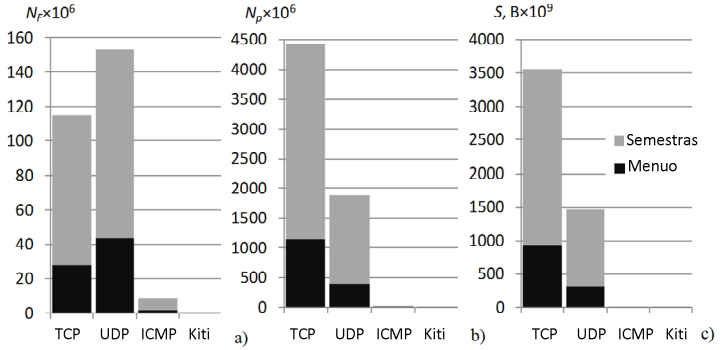
Tinklo srautas yra apibrėžiamas kaip paketų seka, kuri per tam tikrą laiko tarpą yra perduodama tarp dviejų taškų paketų komutavimo tinkle. Duomenų srautai yra aprašomi gavėjo ir siuntėjo IP adresais ir transportinio lygmens prievadų numeriais. *NetFlow* atvaizduoja duomenų šaltinį pakankamai detaliai ir to pakanka plačiai tinklo saugumo analizei.

NetFlow įrašai, gaunami iš Cisco maršrutizatoriaus, pateikia daug informacijos apie srautą tinkle. *NetFlow* 5-oji versija (plačiausiai naudojama) pasižymi šiomis savybėmis (*NetFlow* 2004):

- atvaizduoja sekos numerį;
- identifikuoja įėjimo ir išėjimo prievadus;
- fiksuoja laiko žymes srauto pradžiai ir pabaigai milisekundžių tikslumu;
- pateikia paketų skaičių ir duomenų kiekį sraute;
- atvaizduoja gavėjo ir siuntėjo IP adresus, prievadų numerius, IP protokolo, ir kitų L3/L4 lygmens paketų žymes;
- pateikia kai kurią L3 lygmens maršrutizavimo informaciją.

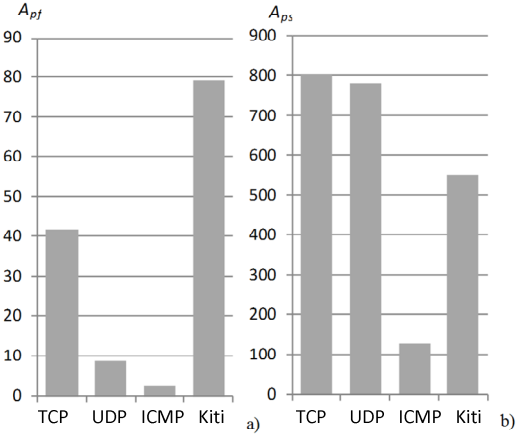
Elektronikos fakulteto kompiuterių tinklas turi 253 viešųjų IP adresų ir apjungia daugiau negu 350 kompiuterių, tame tarpe 5 pašto serverius, 7 web serverius ir 1 FTP serverį.

Šiame skyriuje pateikti duomenys buvo surinkti iš VGTU Elektronikos fakulteto kompiuterių tinklo. Kompiuterių tinklas buvo stebimas visą mokslo metų semestrą (4 mėnesius). Srautų, paketų ir persiųstų duomenų kiekis semestro ir mėnesio bėgyje santykinai yra toks pats (2.2 pav.).



2.2 pav. Tinklo srauto kiekių priklausomybė per mėnesį ir semestrą pagal protokolus: a) srautų kiekis, b) paketų skaičius, c) persiustų duomenų kiekis

Fig. 2.2. Network traffic amount during a semester and a month according to the protocol: a) number of *NetFlow*, b) number of packets and c) the size of transferred data



2.3 pav. Vidutinių reikšmių priklausomybė nuo protokolo: a) paketų skaičius sraute ir b) paketų dydis

Fig. 2.3. Average values: a) number of packets in a *NetFlow* and b) size of the packet, according to the protocol

2.3 paveiksle pavaizduota informacija parodo vidutinį paketų skaičių sraute A_{pf} ir vidutinį paketų dydį A_{ps}, priklausomai nuo protokolo semestro bėgyje. Didžiausias srautų kiekis priklauso UDP protokolui (angl. *User Datagram Protocol*) (2.2 pav.), tačiau TCP (angl. *Transmission Control Protocol*) protokolu buvo persiųstas didžiausias paketų skaičius ir duomenų kiekis. ICMP protokolas (angl. *Internet Control Message Protocol*) yra naudojamas tarnybiniams tikslams, o ne duomenų persiuntimui, todėl šio protokolo paketų skaičius ir duomenų kiekis tinkle yra labai mažas.

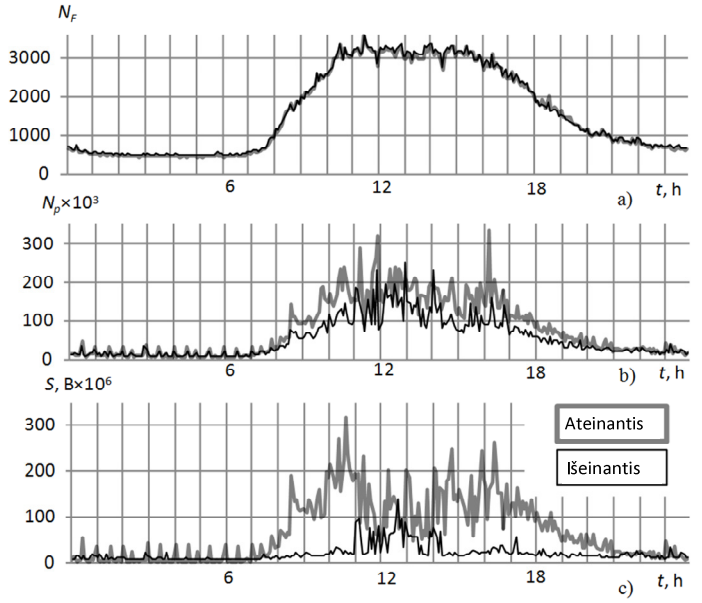
Bendroje statistikoje egzistuoja ir kita grupė paketų. Jai priskiriami įvairūs tuneliavimo ir maršrutizavimo protokolai, kurie nebenaudoja UDP ir TCP transporto protokolų. Remiantis tuo, kad tokių paketų skaičius yra labai mažas (500 kartų mažesnis negu TCP (2.2 pav.)), šis srautas bendros statistikos neįtakoja.

2.2. Srautų kompiuterių tinkle statistinė analizė

Šiame poskyryje pateikta statistikos analizė remiasi vieno mėnesio (2012 metų lapkričio mėn.) surinktais duomenimis. Viso semestro duomenų apdorojimas yra labai komplikuotas, kadangi didelis duomenų kiekis reikalauja daug resursų, bet vieno mėnesio statistika parodo kompiuterių tinklo vartojimo tendenciją per visus metus. *NetFlow* duomenys buvo apdorojami panaudojant *Nfdump* įrankį ir *Perl* skriptus.

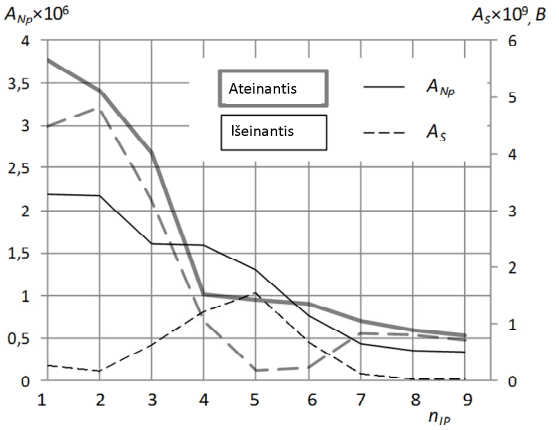
Vidutinė TCP protokolo dienos srauto statistika buvo apskaičiuota remiantis vidutinėmis vienos dienos reikšmėmis mėnesio bėgyje. Srautų skaičius N_F, paketų skaičius N_p ir persiųstų duomenų kiekis S yra pavaizduoti ateinančiam ir išeinančiam srautams (2.4 pav.). Ateinančių ir išeinančių srautų kiekių skirtumas yra minimalus, vidutinis išeinančių srautų kiekis yra 2 % didesnis, ateinančių paketų skaičius yra 30 % didesnis, o duomenų kiekis yra 80 % didesnis.

Statistinė informacija parodo, kad Elektronikos fakulteto tinkle daugiausiai yra tinklo resursų vartotojų (350 kompiuterių tinkle yra naudojami fakulteto darbuotojų ir studentų), bet ne resursų teikėjų (13 serverių dažniausiai yra naudojami vidinėms reikmėms). Tinklo vartojimo statistika pagal unikalų IP adresą parodo didžiausius tinklo srauto vartotojus (2.5 pav.). Vidutinis paketų skaičius A_{Np} ir persiųstų duomenų kiekis A_S geriausiai tai iliustruoja. Paketų skaičiaus vidurkio statistika atvaizduoja rezultatus rūšiuojant pagal unikalų IP adresą. Pirmieji trys unikalūs IP vartotojai vartoja didžiausią įeinančio ir išeinančio srautų dalį lygiomis dalimis, kiti IP adresų vartotojai skirtingomis kryptimis srautą naudoja skirtingai (tas yra matoma ir 2.6 pav., a).



2.4 pav. Vidutinė TCP protokolo dienos srauto statistika: a) srautų skaičius, b) paketų skaičius, c) duomenų kiekis

Fig. 2.4. Average daily TCP traffic: a) number of *NetFlow*, b) number of packets and c) the size of transferred data

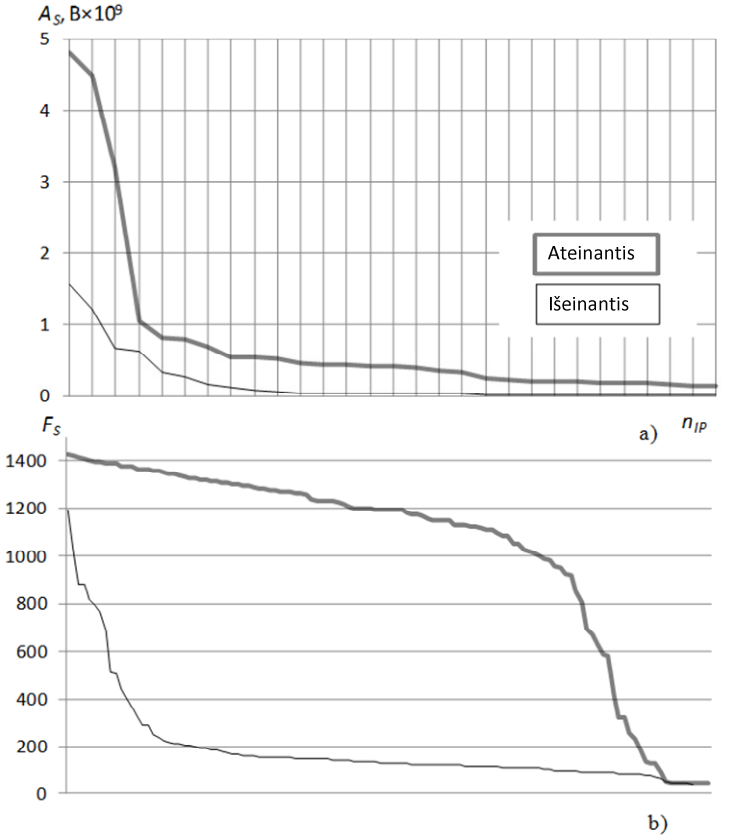


2.5 pav. Paketų skaičiaus vidurkio statistika atvaizduota rezultatus rūšiuojant pagal unikalų IP adresą

Fig. 2.5. Average number of TCP packets and average size of transferred TCP data distribution according to the unique source IP sorted according to packet number

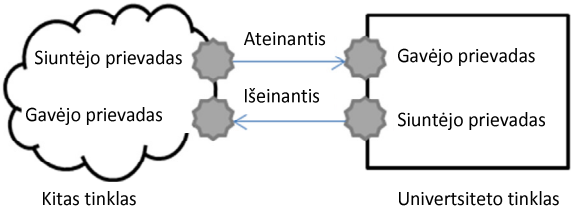
Grafikas parodo, kad didžiausią tinklo srauto dalį generuoja 9 unikalūs IP adresai, o ateinančio srauto didžioji dalis skirta 4-iems IP adresams. Tai parodo, kad tinkle yra naudojamas adresų transliavimas (angl. *Network Adress Translation*).

Persiųstų duomenų vidurkio pasiskirstymas A_S pagal IP adresus yra eksponentės formos (2.6 pav., a), bet jeigu paimti persiųstų paketų kiekio pasiskirstymą, tai ateinančių ir išeinančių srautų pasiskirstymai skiriasi (2.6 pav., b).



2.6 pav. Vidutinis persiųstų TCP duomenų kiekio pasiskirstymas pagal unikalius IP adresus a) ir paketų dydžių pasiskirstymas b)
Fig. 2.6. Average values of a) size of transferred TCP data distribution according to the unique IP and b) packet size distribution

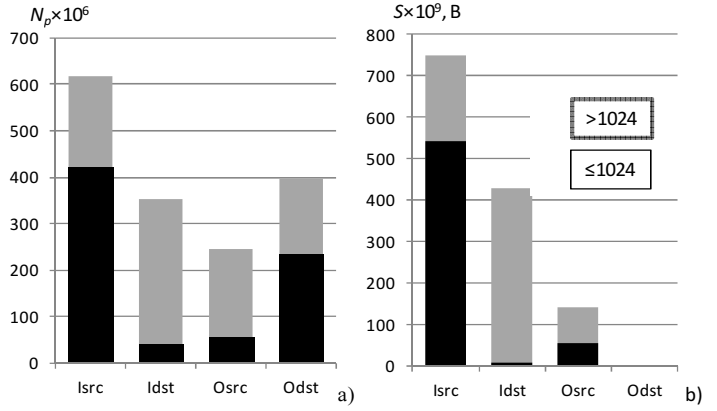
NetFlows protokolas kaupia informaciją apie ateinančius ir išeinančius srautus (2.7 pav.) remiantis šaltinio ir gavėjo transporto lygmens prievadais.



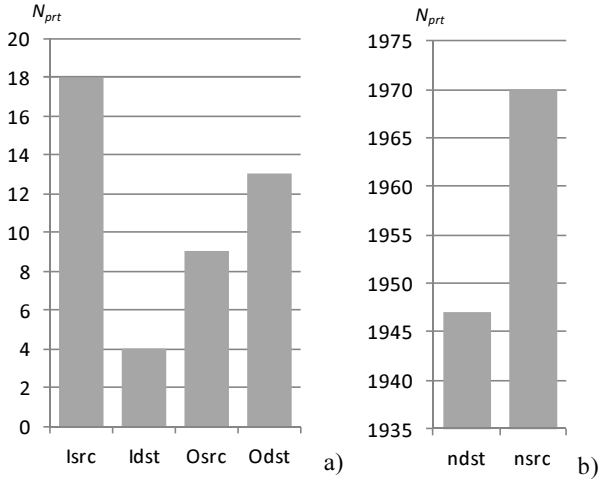
2.7 pav. Komunikavimas tarp tiriamojo kompiuterio tinklo ir kitų tinklų
Fig. 2.7. Communication between the computer network which is an object of the research and other networks

TCP komunikacijos metu vartotojas siunčia užklausą dėl informacijos gavimo iš serverio, komunikacijai yra naudojami transporto lygmens vartotojo šaltinio ir serverio gavėjo prievadai. Gavėjo prievadas dažniausiai yra susijęs su serviso (paslaugos) prievadu, kurio rėžiai priimta laikyti yra iki 1 024 – tai yra gerai žinomi tinklo paslaugų prievadai.

Prievadai, kurių skaičius yra didesnis už 1 024 dažniausiai yra naudojami kaip šaltinio prievadai naujoms ir nestandartinėms paslaugoms. Šuo metu servisų skaičius tinkle yra labai išaugęs, dėl to paminėta taisyklė nėra visada taikytina, bet informacija apie gerai žinomas paslaugas, kurioms galioja ši taisyklė, yra labai svarbi (2.8 pav.). Grafike parodyta statistika, kuomet siuntėjo prievadas yra didesnis už gavėjo prievadą. Labiausiai paplitusi yra HTTP paslauga.

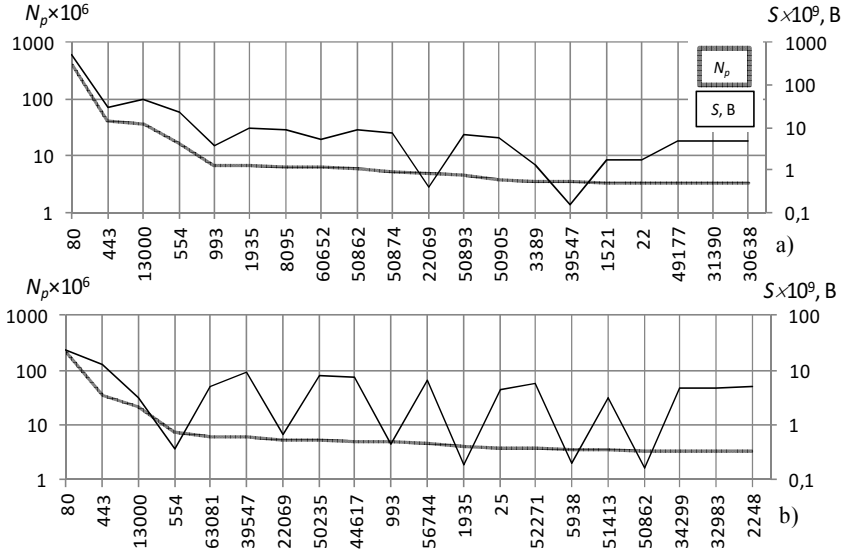


2.8 pav. Vidutinio ateinančio ir išeinančio srautų skaičiaus priklausomybė nuo gavėjo ir siuntėjo prievado numerio >1 024 arba <1 024: a) paketai, b) baitai
Fig. 2.8. Average number of incoming and outgoing traffic according to source and destination ports evaluated in: a) packets and b) bytes



2.9 pav. Prievadų pasiskirstymas pagal: a) gerai žinomus <1 024 prievadus ir b) bendras prievadų skaičius

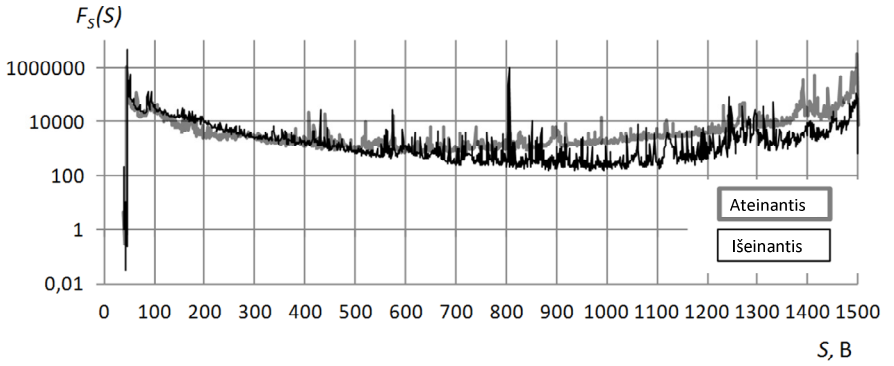
Fig. 2.9. Distribution of ports according to a) incoming and outgoing traffic well-known ports and b) overall number of destination and source ports



2.10 pav. Prievadų pasiskirstymas pagal: a) gerai žinomus <1 024 prievadus ir b) bendras prievadų skaičius

Fig. 2.10. Average number of a) incoming and b) outgoing traffic according to number of packets and size of transferred data in bytes per port

TCP komunikacijoje yra naudojami transporto lygmens prievadai, kurie gali būti suskirstyti į grupes: sistemų prievadai (0–1 023), vartotojo prievadai (1 024–49 151), dinaminiai arba/ir privatūs prievadai (49 152–65 535) (Service Name and Transport Protocol Port Number Registry 2014). Ateinančio srauto šaltinio prievado numeris yra didesnis už gavėjui skirtą serviso prievado numerį (2.9 pav., a). Tinklo srauto dydžio pasiskirstymas yra eksponentinis, kai rezultatas yra surūšiuotas pagal paketų skaičių (2.10 pav.).



2.11 pav. Paketų dydžio pasiskirstymas
Fig. 2.11. Packet size distribution

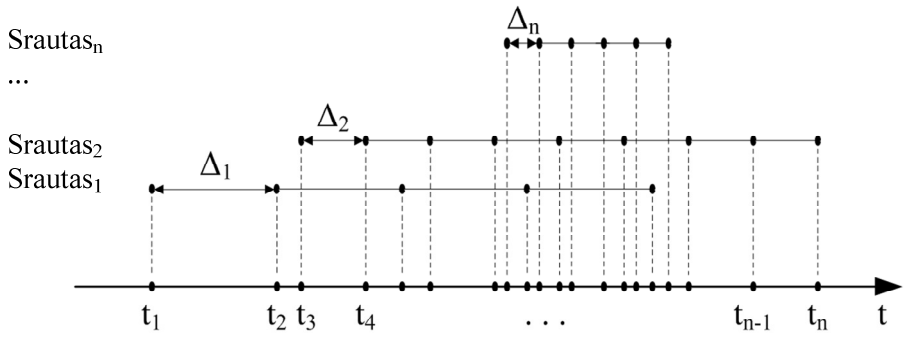
Ateinantieji ir išeinantieji paketai turi panašią formą, bet ateinantieji paketai yra linkę turėti didesnį paketo dydį (2.11 pav.). Aukščiau pateiktas pasiskirstymas parodo, kad yra tam tikrų servisų, kurie nepatenka į bendrą paketų dydžio pasiskirstymo tendenciją: pikas virš paketo dydžio 800 B yra vienos dienos aktyvumas, kai iš FTP serverio buvo siunčiama 36 GB duomenų, o paketų dydžio pasiskirstymas buvo: 804, 805, 806 ir 808 B.

2.3. Atėjimo laiko tarp paketų pasiskirstymo tyrimas

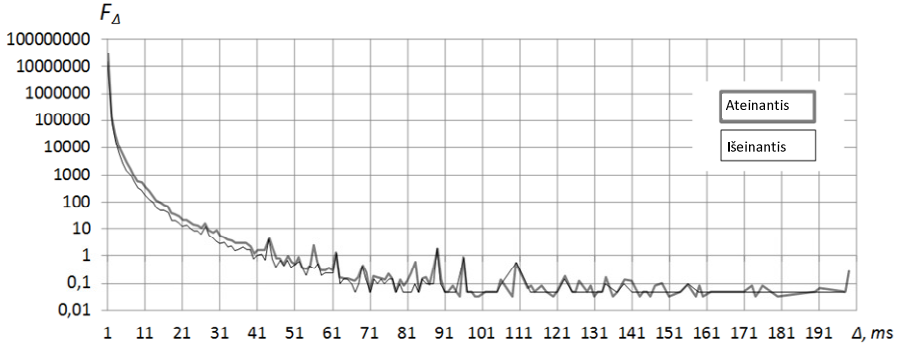
Tinklo laikinės charakteristikos yra labai svarbios norint modeliuoti tinklą. Paketų atėjimo laiko pasiskirstymas buvo skaičiuojamas pasinaudojant (2.12 pav.) pateikta logika.

Buvo priimta, kad visi paketai sraute yra pasiskirstę tolygiai. Sraute *Flow1* laikas tarp paketų yra lygus Δ_1, sraute *Flow2* – Δ_2, o sraute *Flown* – Δ_n. Tačiau realiame tinkle suminis paketų atėjimo laikas yra pasiskirstęs netolygiai. Kaip žinome, paketai tinkle yra siunčiami vienas paskui kitą. Po paketų atėjimo laiko $t_1, t_2, \ldots, t_n$ duomenų apdorojimo, kiekvienam srautui gaunama, kad laikas tarp ateinančių paketų tinkle yra atsitiktinai pasiskirstęs. Iš teorijos žinoma, kad jeigu

sumuojasi keli srautai su įvairiais pasiskirstymo dėsniais, tuomet suminio srauto pasiskirstymas gaunamas artimas rodikliniam dėsniui. Be to, kuo daugiau tokių srautų sumuojasi, tuo suminio srauto pasiskirstymas artimesnis rodikliniam pasiskirstymo dėsniui. Iš paketų pasiskirstymo tinkle galima apskaičiuoti paketų atėjimo laiko pasiskirstymą. Gautas pasiskirstymas yra parodytas 2.13 paveiksle.



2.12 pav. Paketų pasiskirstymas tinklo sraute
Fig. 2.12. Assumed distribution of packets in network flows



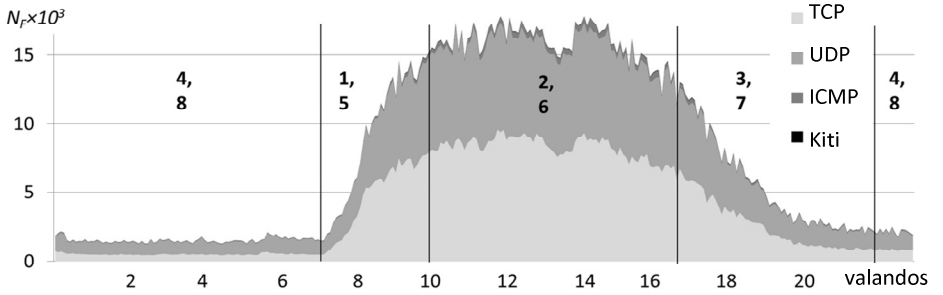
2.13 pav. Atėjimo laiko tarp paketų pasiskirstymas
Fig. 2.13. Distribution of time interval between the packets

Paketų atėjimo laiko pasiskirstymas parodytas 2.13 paveiksle yra tokios pat formos abejomis tinklo kryptimis. Kaip matome, jis yra artimas rodikliniam pasiskirstymo dėsniui. Akademinis kompiuterių tinklas nėra apkrautas, todėl didžiausia dalis paketų yra pasiskirsčiusi su mažais intervalais tarp jų.

Tinklo srautas dienos laike buvo padalintas į sekcijas pagal kryptį ir apkrovą dienos metu, vėliau buvo aprašyti atėjimo laiko tarp paketų pasiskirstymai. Buvo atliktas Kolmogorovo-Smirnovo testas ir išsiaiškinta, kuris pasiskirstymo dėsnis

labiausiai tinka kiekvienai sekcijai. Kompiuterių tinklo modeliavimas reikalauja kompiuterių tinklo charakteristikų ir pasiskirstymo dėsnių žinojimo. Pasiskirstymo dėsnio žinojimas palengvina darbą tinklo modeliavimo metu.

NetFlow duomenys buvo surinkti 2012 metų spalio mėnesį. Analizei buvo naudojama tik darbo dienų statistika, kadangi visų savaitgalių srautas sudaro vos 7,3 % viso mėnesio srauto. Šį analizė yra skirta paketų atėjimo laiko pasiskirstymo nustatymui, dėl to paketų dydžiai, siuntėjo ir gavėjo prievadai bei IP adresai nėra įvertinti. Vidutinis dienos srautas buvo apskaičiuotas atsižvelgiant į visų mėnesio darbo dienų vidutinius atėjimo ir išėjimo srautus. Srautų kiekis N_F dienos bėgyje yra parodytas 2.14 paveiksle.



2.14 pav. Vidutinis tinklo srautų kiekis dienos bėgyje: TCP, UDP, ICMP ir kitiems protokolams

Fig. 2.14. Average daily network traffic in number of flows: TCP, UDP, ICMP and Other

Tinklo srautų grafikas yra padalintas į keturias sekcijas priklausomai nuo tinklo apkrovos: tinklas nėra naudojamas naktį, tik tam tikra tarnybinė informacija yra persiunčiama 22.00–7.00 val. (4, 8); tinklas labiausiai yra naudojamas dienos metu 10.00–16.30 val. (2, 7); tinklo naudojimo pradžia 7.00–10.00 val.(1, 5) ir pabaiga 16.30–22.00 val. (3, 7), kai vartotojai ateina ir išeina iš fakulteto. Tinklo srautas gali būti apibūdintas jo kryptimi: ateinantis srautas (1–4), kuris ateina į fakulteto tinklą ir išeinantis (5–8), kuris išeina iš fakulteto tinklo. Dienos sekcijos yra aprašytos 2.1 lentelėje.

Tinklo srauto sekcijos yra skirtingos trukmės – piko metu sekcijos trukmė yra antra pagal ilgį (2, 6), bet sudaro didžiąją srauto dalį, o nakties periodo sekcija (4, 8) yra ilgiausia, bet sudaro mažiausią srauto dalį. Tinklo srauto apkrovos augimas (1, 5) yra staigesnis negu sumažėjimas (3, 7). Tinklo srauto statistika pagal transporto protokolą yra parodyta 2.2 ir 2.3 lentelėse atitinkamai ateinančiam ir išeinančiam srautams.

2.1 lentelė. Tinklo srautų sekcijos dienos metu
Table 2.1. Network Traffic Sections

Ateinantis srautas	Paketai, %	Laikas	Trukmė, h	Paketai, %	Išeinantis srautas
1	6,96	7.00–10.00	3	4,36	5
2	36,58	10.00–16.30	6,30	23,05	6
3	16,48	16.30–22.00	5,30	10,78	7
4	0,79	22.00–7.00	9	1,01	8

2.2 lentelė. Ateinančio tinklo srauto statistika
Table 2.2. Incoming Network Traffic Statistics

Parametrai \ Protokolas	TCP	UDP	ICMP	Kiti
Srautas, B×10^9	843	257	0,158	36
Srautas, vnt.	12 954 137	11 733 655	438 532	87 850
Paketai, ×10^6	1 007	265	2	30
Vidutinis paketų skaičius sraute	78	23	4	339
Vidutinis paketų dydis, B	838	970	96	1 197

Duomenys 2.2 ir 2.3 lentelėse neatvaizduoja 6,4 % paketų, kurie patenka tarp duomenų nuskaitymo ribų, t. y. *nfdump* priemonė negali apdoroti srautų, kurie prasideda ir baigiasi skirtinguose duomenų nuskaitymo perioduose.

2.3 lentelė. Išeinančio tinklo srauto statistika
Table 2.3. Outgoing Network Traffic Statistics

Parametrai \ Protokolas	TCP	UDP	ICMP	Kiti
Srautas, B×10^9	131	91	0.358	3
Srautas, vnt.	12 072 179	10 958 168	882 523	35 742
Paketai, ×10^6	464	169	3	14
Vidutinis paketų skaičius sraute	38	15	3	384
Vidutinis paketų dydis, B	282	539	125	228

Ateinantis srautas dominuoja, jame TCP yra 6,4 karto daugiau ir 2,8 karto daugiau UDP ateinančių duomenų, tačiau ateinančių duomenų srautų skaičius yra tik 7 % didesnis abiem protokolams, o paketų skaičius yra 2,2 karto didesnis TCP protokolui ir 1,6 karto didesnis UDP protokolui už išeinančių duomenų srautus atitinkamiems protokolams. Vidutinis paketų skaičius sraute yra didesnis

ateinančiam srautui ir tai galioja abiem protokolams, bet yra skirtumas pagal paketų dydį: ateinantys TCP paketai yra 3 kartus didesni, o UDP paketai – 1,8 karto didesni už išeinančio srauto paketus.

Galima pastebėti, kad TCP protokolu persiųstų duomenų kiekis yra 2,8 karto didesnis už UDP protokolu persiųstų duomenų kiekį, bet srautų kiekis yra vos 1 % didesnis už UDP srautų kiekį. Vidutinis paketų skaičius TCP protokolui yra 3 kartus didesnis. UDP protokolo paketų vidutinis paketo dydis yra 26 % didesnis už TCP. Duomenų apimtis, paketų ir srautų skaičius tokių protokolų kaip ICMP ir kitų, yra labai maži, dėl to jie nėra nagrinėjami.

Akademinis kompiuterių tinklas yra sujungtas panaudojant Ethernet technologiją su 100 Mbps ir 1 Gbps tinklo segmentais. 100 Mbps Ethernet turi 0,96 µs minimalų laiką tarp paketų, o 1 Gbps Ethernet tinklui minimalus laikas tarp paketų yra lygus 0,096 µs. Paketų dydžiai tinkle svyruoja tarp 64 ir 1 518 baitų. Dėl to 1 Gbps tinklo segmente minimalus atėjimo laikas tarp paketų, svyruoja tarp 0,608 µs ir 12,240 µs. Atitinkamai 100 Mbps tinklo segmentui minimalus laikas tarp ateinančių paketų svyruoja nuo 6,080 µs iki 122,400 µs. Laiko tikslumas žemiau pateiktuose skaičiavimuose buvo pasirinktas 0,1 ms žinant, kad vidutinis paketų atėjimo laikas tinkle yra 2,835 ms. Buvo įvertinta, kad 1 ms laiko tikslumas nebus pakankamas, kadangi į pirmą 1 ms intervalą patektų maždaug 80 % visų paketų. Tikslumo didesnio negu 0,1 ms panaudojimas yra neįmanomas, dėl to kad *NetFlow* protokolas sugeba įrašyti duomenis tik 1 ms tikslumu, bet 0,1 ms laiko tikslumas yra galimas, kadangi viename sraute būna daugiau negu vienas paketas.

2.4. Pasiskirstymo matematinė išraiška

Akademinis kompiuterių tinklas nėra apkrautas, o tinklo resursai yra naudojami kai jų reikia. Dėl netolygaus tinklo panaudojimo ir galimų srauto panaudojimo pikų, kurie gali įtakoti bendrą rezultatą, buvo įvestos išskirtys. Išskirtys yra naudojamos tam, kad eliminuoti tokią situaciją, kai pavieniai tinklo srauto pikai įtakoją visą tinklo apkrovimą. Neįvedus išskirčių galima tokia situacija, kai visos dienos paketų skaičius yra nedidelis, pvz. 10^5 paketų eilės, tuomet vienos dienos bėgyje atsiradęs 10^3 paketų pikas jau įtakotų bendrą rezultatą. Pagal apibrėžimą išskirtys – tai netipinės ir retos reikšmės, kurios yra žymiai nukrypusios nuo kitų duomenų pasiskirstymo. Šie duomenys arba tikrai atspindi tiriamojo reiškinio (kintamojo) tikrąsias savybes, arba yra matavimo klaidų, arba anomalinių reiškinių padarinys. Dėl to išskirtys dažnai neįtraukiamos į tiriamąjį modelį.

Sąlygine išskirtimi vadinamas duomuo priklausantis intervalui žemiau pateiktoje formulėje (2.1 formulė):

$$[Q1 - 3 \cdot IQR,\ Q3 + 3 \cdot IQR],\ \ IQR = Q3 - Q1. \qquad (2.1)$$

Tinklo paketų atėjimo laiko reikšmės buvo padalintos į rėžius: $((n-1)\times 10^{-4}, n\times 10^{-4}]$, $n \in [1, 5\,000]$.

2.4 lentelė. Išskirčių dalis visame sraute
Table 2.4. Outlier percentage

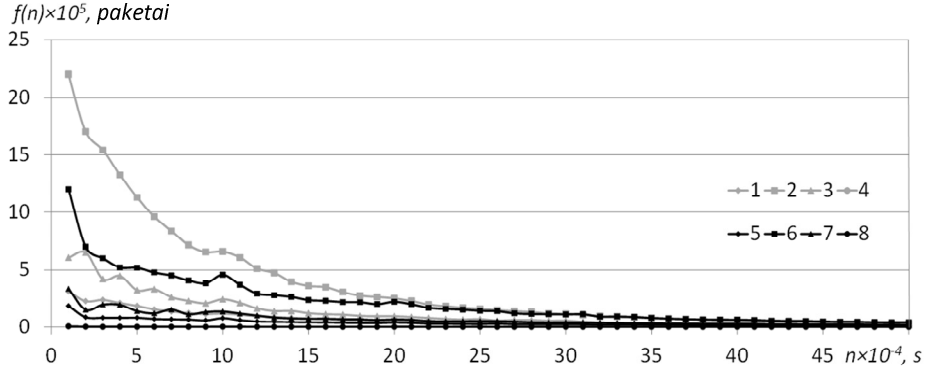
Dienos dalis, bendras srautas	TCP protokolas		UDP protokolas	
	Išskirtys, %	Paketai, %	Išskirtys, %	Paketai, %
1	3,80	5,07	4,22	56,09
2	1,67	2,26	5,20	52,04
3	4,58	7,56	5,37	21,58
4	3,61	47,91	3,62	15,92
5	4,17	6,95	2,13	36,66
6	1,43	4,44	2,27	11,53
7	4,40	12,16	2,63	34,84
8	3,47	56,12	4,71	30,98
Ateinantis srautas	3,41	15,70	4,60	36,41
Išeinantis srautas	3,37	19,92	2,94	28,50
Vidurkis	3,39	17,81	3,77	32,46

Didesnis rėžių intervalas nėra naudojamas dėl mažo paketų skaičiaus patenkančio į jį. Pateikti grafikai (2.15, 2.16, 2.17, 2.18 pav.) atvaizduoja tik pasirinkto rėžio pasiskirstymą, atmetant ilgą grafiko uodegą. Suminis ateinančių paketų laiko reikšmių skaičius vienam transporto protokolui yra 23 (darbo dienos) × 5 000 (rėžių skaičius) = 115 000 (laiko reikšmių). Išskirtys, tai yra tam tikro laiko tam tikros dienos duomenys, kurie nepateko į atvaizduojamą rezultatą ir buvo atmesti. Informacija (procentais) apie išmestus dėl išskirčių paketus ir paketus, įtrauktus į skaičiavimą yra pateikta 2.4 lentelėje.

3,39 % išskirčių buvo išmesta iš TCP protokolo ir 3,77 % iš UDP protokolo srautų. Tai sudaro apie 17,81 % TCP ir 32,46 % UDP atmestų ir nepanaudotų skaičiavimams paketų. Išeinantiems srautams buvo mažiau išskirčių: TCP srautui apie 0,04 % ir UDP – 1,66 %. Didžiausias išskirčių skaičius TCP srautui yra dienos 3 sekcijoje – 4,58 % (7,56 % visų paketų) ir UDP srautui dienos 8 sekcijoje – 4,71 % (30,98 % visų paketų).

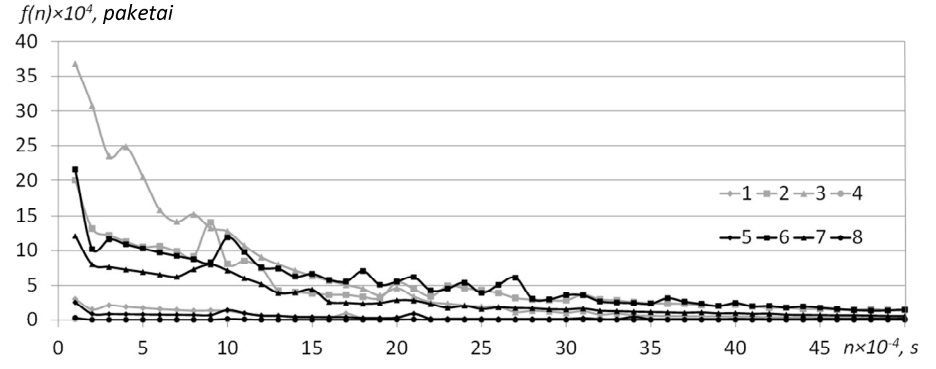
Silpnai apkrautas kompiuterių tinklas turi didelį išskirčių skaičių dėl srauto išnaudojimo pikų. Labiau apkrautame kompiuterių tinkle išskirčių skaičius mažėja. UDP sraute išskirčių skaičius yra didesnis negu TCP sraute.

Paketų atėjimo laiko tarp paketų pasiskirstymas *f(n)* pagal srauto sekcijas yra parodytas 2.15 pav. ir 2.16 pav., kur *n* parodo rėžio pabaigą.



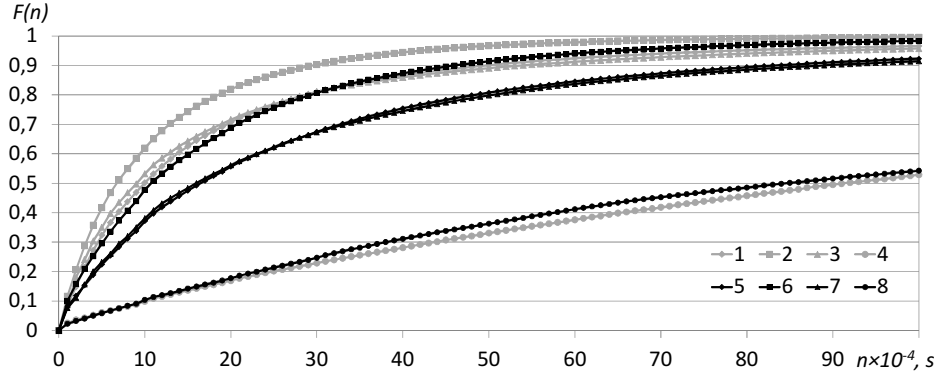
2.15 pav. Paketų atėjimo laikų pasiskirstymas TCP srautui
Fig. 2.15. Distribution of TCP packet inter-arrival time

Didžiausios reikšmės yra 2 ir 6 sekcijose, kur yra srauto pikai, tuo tarpu mažiausios reikšmės yra 4 ir 8 sekcijose, kurios atvaizduoja nakties periodą.



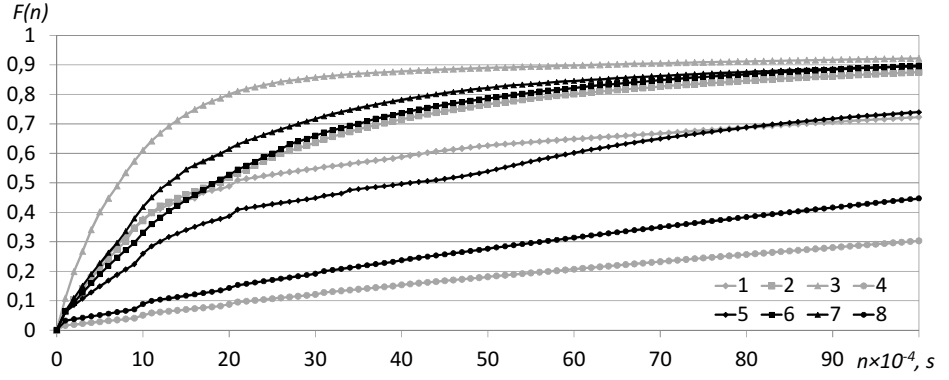
2.16 pav. Paketų atėjimo laikų pasiskirstymas UDP srautui
Fig. 2.16. Distribution of UDP packet inter-arrival time

Paketų atėjimo laiko pasiskirstymas tiesiogiai priklauso nuo paketų skaičiaus, o kaupiamoji pasiskirstymo funkcija *F(n)* normalizuoja grafiką ir leidžia palyginti tendencijas nepriklausomai nuo paketų skaičiaus. Grafikai TCP ir UDP srautams yra parodyti 2.17 ir 2.18 paveiksluose.



2.17 pav. Normuotas paketų atėjimo laikų pasiskirstymo grafikas TCP srautui
Fig. 2.17. CDF of TCP packet inter-arrival time

Normuotas TCP srauto grafikas parodo, kad 70 % visų TCP paketų patenka į 1, 3 ir 6 sekcijas, kai paketų atėjimo laikas yra trumpesnis negu 2 ms ir 80 % visų paketų patenka į 2 sekciją, kai tinklo apkrovimas yra didelis. 50 % UDP srauto patenka į 1, 2 ir 6 sekcijas, 60 % – į 7 sekciją, 80 % – į 3 sekciją, kai paketų atėjimo laikas yra mažesnis negu 2 ms (2.18 pav.).



2.18 pav. Normuotas paketų atėjimo laikų pasiskirstymo grafikas UDP srautui
Fig. 2.18 CDF of UDP packet inter-arrival time

Kompiuterių tinklo srautų modeliavimui reikalingos jų matematinės išraiškos. Eksperimentiškai gautų paketų atėjimo laikų matematiniam aprašymui buvo panaudotas Kolmogorov-Smirnov kriterijus, kurio pagalba buvo rastas teorinis skirstinys labiausiai atitinkantis empirinį. Palyginimui buvo panaudoti Veibulo, Pareto, Gama, Eksponentinis ir Normalus pasiskirstymai. Suderinamumo tikrinimo testas buvo atliekamas su paketų atėjimo laiko

pasiskirstymu, remiantis vidutiniu abiejų krypčių tinklo srautu (žr. 2.5 lentelę). Veibulo, Pareto, Normalus ir Gama pasiskirstymai sudaromi panaudojant formos ir skalės parametrus. Eksponentinis pasiskirstymas sudaromas panaudojant dažnio parametrą. Formos parametras α paveikia grafiko formą, o skalės parametras β ištempia arba sutraukia grafiką. Pirmas formos parametras 2.5 lentelėje yra atvaizduojamas kaip Param. 1, o antras skalės parametras – Param. 2. Kolmogorov-Smirnov parametras *KS* parodo didžiausio absoliutaus skirtumo reikšmę tarp eksperimentinės ir teorinės kreivių, mažesnis skaičius išreiškia didesnį atitikimą. Daugiklis *A* yra naudojamas teorinės kreivės priderinimui prie eksperimentinės kreivės *y* ašyje.

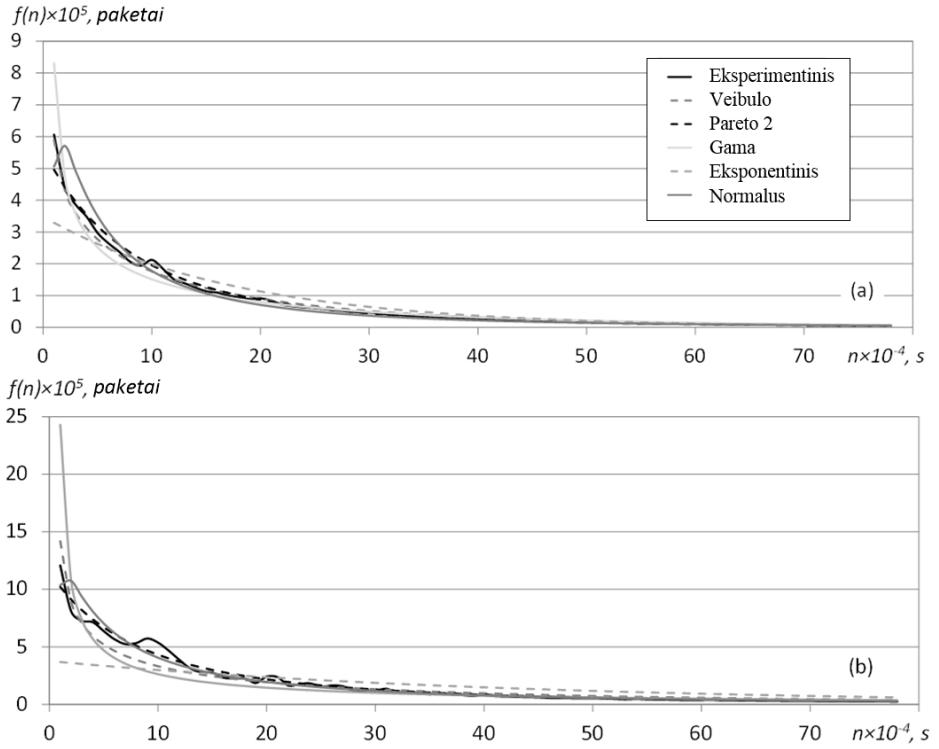
2.5 lentelė. Suderinamumo tikrinimo testas paketų atėjimo laiko pasiskirstymui
Table 2.5. Goodness-of-fit of packet inter-arrival time distribution

Protokolas	Pasiskirstymas	Param. 1	Param. 2	*KS*	*A*	Vieta
TCP	Veibulo	0,798	0,002	0,063	609,64	2
	Pareto 2	2,658	0,003	0,058	606,03	1
	Gama	0,509	0,004	0,129	638,19	-
	Eksponentinis	561,680	-	0,123	603,03	-
	Normalus	1,409	-7,151	0,075	602,14	3
UDP	Veibulo	0,734	0,003	0,053	322,35	-
	Pareto 2	1,848	0,004	0,046	322,90	1
	Gama	0,480	0,008	0,103	333,79	-
	Eksponentinis	249,330	-	0,176	317,91	-
	Normalus	1,545	-6,472	0,053	323,57	2

Kaip matome iš 2.5 lentelės, kompiuterių tinklo srauto vidurkiui Pareto du (Pareto 2) pasiskirstymas yra labiausiai panašus į eksperimentinius pasiskirstymus abiem protokolų srautams TCP ir UDP (2.19 pav.).

Pareto 2 pasiskirstymas yra standartinis Pareto pasiskirstymas su pasislinkusia x ašimi, o grafikas patenka į intervalą $0 \le x < +\infty$, kai standartinis Pareto pasiskirstymas patenka į $\beta \le x < +\infty$ intervalą. Pareto 2 statistinis pasiskirstymas dar gali būti pavadintas kaip Lomax pasiskirstymas ir dažniausiai yra naudojamas verslo ir ekonomikos uždavinių modeliavime.

Tinklo srauto sekcijos buvo įvertintos ir suderinamumo tikrinimo testas buvo atliekamas tam, kad išsiaiškinti ar yra paketų atėjimo laiko pasiskirstymo priklausomybė nuo dienos laiko.



2.19 pav. Suderinamumo tikrinimo testas paketų atėjimo laiko pasiskirstymui: (a) TCP, (b) UDP srautams

Fig. 2.19 Goodness-of-fit of packet inter-arrival time distributions: (a) TCP, (b) UDP

2.6 lentelė. TCP paketų atėjimo laikų pasiskirstymo funkcijų koeficientai

Table 2.6. TCP packet inter-arrival time distribution coefficients

Protokolas	Sekcija	Pasiskirstymas	α	β	KS	A
TCP	1	Pareto2	2,7278	0,0034	0,0478	359,63
	2	Veibulo	0,8978	0,0011	0,0623	1890,70
	3	Pareto2	1,9981	0,0022	0,0558	698,54
	4	Pareto2	3,1931	0,0361	0,0246	23,18
	5	Veibulo	0,7750	0,0027	0,0441	230,12
	6	Gama	0,7810	0,0023	0,0545	1201,46
	7	Veibulo	0,7578	0,0028	0,0467	437,86
	8	Veibulo	0,7897	0,014	0,0211	24,93

2.7 lentelė. UDP paketų atėjimo laikų pasiskirstymo funkcijų koeficientai
Table 2.7. UDP packet inter-arrival time distribution coefficients

Protokolas	Sekcija	Pasiskirstymas	α	β	KS	A
UDP	1	Normalus	1,9497	-6,084	0,0560	46,55
	2	Pareto2	1,8484	0,0039	0,0456	322,91
	3	Pareto2	1,2351	0,0009	0,0643	339,50
	4	Pareto2	4,0406	0,1065	0,0221	10,35
	5	Veibulo	0,6066	0,0073	0,0476	40,96
	6	Pareto2	1,9606	0,0042	0,0409	338,45
	7	Pareto2	1,4612	0,0022	0,0454	184,25
	8	Pareto2	2,7942	0,0386	0,0293	10,15

Visiems pasiskirstymo TCP srautų sekcijų grafikams buvo rastas teorinis pasiskirstymas, UDP srauto pritaikymas buvo komplikuotas (2.6 ir 2.7 lenteles): 1 ir 5 kreivės nesutapo su eksperimentinėmis kreivėmis.

Pareto 2 tinka didžiajai daliai kreivių, o būtent ten kur tinklo srautas auga ir mažėja. Pareto 2 pasiskirstymas yra labiausiai (arba antras iš eilės) tinkamas pasiskirstymas kreivėms apibūdinti, dėl to Patreto 2 pasiskirstymo dėsnis gali būti taikomas tinklo modelyje, kaip paketų atėjimo laiko pasiskirstymo funkcija.

2.5. Antrojo skyriaus išvados

1. Kompiuterių tinkle *NetFlows* protokolu surinktų duomenų analizė parodo, kad paketų atėjimo laiko pasiskirstymas nepriklauso nuo paketų srauto krypties.

2. Tinklo srauto padalinimas į sekcijas leidžia nustatyti vyraujančias tendencijas tinkle, kurios reikalingos tiksliam tinklo modelio sudarymui. Tai leidžia sumažinti analizuojamų duomenų kiekį ir tiksliau nustatyti to paros laikotarpio tinklo srauto tendenciją.

3. Analizuojant gautus duomenis buvo įvertintos išskirtys, kurios sudaro apie 3,6 % viso TCP ir UDP srauto. Paketų, patenkančių į išskirčių dalį, skaičius priklauso nuo dienos laiko ir tinklo panaudojimo – daugiausiai išskirčių yra nakties metu, kai tinklas mažai naudojamas.

4. Kolmogorov-Smirnov suderinamumo tikrinimo testas parodo, kad TCP ir UDP tinklo srauto paketų atėjimo laiko pasiskirstymo dėsnį tiksliau apibūdinti leidžia Pareto 2 pasiskirstymas negu Puasono. Eksperimento rezultatai parodo, kad TCP ir UDP tinklo srauto pasiskirstymo kreivės yra tokios pat formos.

3

Informacinės sistemos išliekamumo įvertinimas remiantis saugumo analize

Informacinės sistemos saugumo įvertinimui reikalingos žinios apie sistemą ir jai kylančias grėsmes. Sistemos saugumui, kuris tiesiogiai priklauso nuo saugumo mechanizmų, įvertinti patogu naudoti išliekamumo charakteristiką ir pagal ją nustatyti sistemos saugumo pokytį kintant ją veikiantiems veiksniams ir jos parametrams. Išliekamumas – tai sistemos gebėjimas atlikti savo misiją, net jei ji yra puolama ar dalis jos nereikšmingų paslaugų yra sukompromituota (Garšva *et al.* 2006). Taigi išliekamumas – tai sistemos gebėjimas teikti kritines paslaugas net sėkmingo įsibrovimo atveju ir sugebėti atsistatyti iki normalios būsenos atakai pasibaigus. Todėl reikia, kad sistema sugebėtų reaguoti į įvykius ir įvertinti juos: atpažintų atakas, būtų atspari vienoms ir sugebėtų atsikurti po kitų poveikio. Sistemos išliekamumas apima tokias sritis kaip patikimumas, saugumas ir atsparumas trikdžiams. Ši informacija gali būti surinkta remiantis rizikos analize.

Sistemos išliekamumas yra viena iš svarbiausių skaitinių charakteristikų sistemos būsenai esant incidentui įvertinti. Ši charakteristika yra naudojama sistemų palyginimui ir saugumo mechanizmų įvertinimui. Šiame darbe yra siūlomas informacinių sistemų išliekamumo stochastinis modelis. Modelio

parametrai yra paimti iš Lietuvos Respublikos Vidaus reikalų ministro įsakymu „Dėl valstybės institucijų ir įstaigų informacinių sistemų elektroninės informacijos techninių saugos reikalavimų“ teisės akto Nr. 1V-384 patvirtinto 2008 m. spalio 27 d. Šio skyriaus medžiaga paskelbta dviejuose straipsniuose (Garšva, Paulauskas, Gulbinovič, Stankevičius, 2011; Paulauskas, Garšva, Gulbinovič, Stankevičius, Poviliauskas, 2012).

2013 m. spalio 4 d. Lietuvos Respublikos vidaus reikalų ministro įsakymu Nr. 1V-832 aukščiau minimas teisės aktas buvo pripažintas netekusiu galios ir patvirtinti nauji „Techniniai valstybės registrų (kadastrų), žinybinių registrų, valstybės informacinių sistemų ir kitų informacinių sistemų elektroninės informacijos saugos reikalavimai“, kuriuose nustatomi minimalūs elektroninės informacijos saugos techniniai reikalavimai valstybės registrams (kadastrams), žinybiniams registrams, valstybės informacinėms sistemoms ir kitoms informacinėms sistemoms. Šiame darbe toliau yra nagrinėjamas 2008 metais patvirtinti reikalavimai, kadangi tyrimo atlikimo metu naujas teisės aktas dar nebuvo priimtas.

3.1. Saugumą reguliuojantis įstatymas

Vyriausybės informacinės sistemos yra vienos geriausiai valdomų sistemų, todėl jos buvo naudojamos atliekant tyrimą. Informacinės sistemos yra padalintos į kategorijas pagal svarbą (Valstybės institucijų ir įstaigų informacinių sistemų klasifikavimo pagal jose tvarkomą elektroninę informaciją gairės, 2007). Kiekvienai informacinės sistemos kategorijai keliami skirtingi reikalavimai sistemos pasiekiamumui ir atstatymo laikui (Valstybės institucijų ir įstaigų informacinių sistemų elektroninės informacijos techniniai saugos reikalavimai, 2008). Sistemos skirstomos į keturias kategorijas. Reikalavimai pirmos ir antros kategorijos sistemoms yra labai aukšti, sistemos atstatymo laikas neturi būti ilgesnis nei 15 min. pirmajai ir ne ilgesnis negu 1 val. antrajai kategorijai. Reikalavimai trečiosios ir ketvirtosios kategorijos sistemoms yra nustatyti tik darbo dienoms ir darbo valandoms, atitinkamai 8 val. – trečiajai ir 16 val. – ketvirtajai kategorijai. Informacinės sistemos prieinamumas ketvirtosios kategorijos informacinėms sistemoms turi būti užtikrintas ne mažiau kaip 70 % laiko darbo metu darbo dienomis, trečiosios kategorijos informacinėms sistemoms – ne mažiau kaip 90 % laiko darbo metu darbo valandomis, antrosios kategorijos informacinėms sistemoms – ne mažiau kaip 96 % laiko visą parą, pirmosios kategorijos informacinėms sistemoms – ne mažiau kaip 99 % laiko visą parą. Modeliavimui buvo pasirinkta trečiosios kategorijos informacinė sistema, kadangi dauguma vyriausybės informacinių sistemų priklauso trečiajai kategorijai. Trečiosios kategorijos informacinės sistemos darbas turi būti

atstatytas per 8 val., o informacija pasiekiama 90 % laiko darbo valandomis. Informacinių sistemų kategorijų apibendrinimas yra pateikiamas 3.1 lentelėje.

Saugumo mechanizmai yra aprašyti kiekvienai kategorijai. Adaptuojant saugumą reguliuojantį įstatymą prie modeliuojamos aplinkos buvo apibrėžti 36 skirtingi saugumą užtikrinantys mechanizmai. Vienas mechanizmas gali būti naudojamas apsaugai prieš vieną arba kelias grėsmes.

3.1 lentelė. Lietuvos valstybinių institucijų sistemų kategorijos
Table 3.1. Requirements to the Lithuanian government system accessibility and recovery time

Kategorija	Sistemos atstatymo laikas	Informacijos pasiekiamumas	Modulių skaičius
I	15 min.	99 %	7
II	1 val.	96 %	5
III	8 d. d. val.	90 % d. d.	3
IV	16 d. d. val.	70 % d. d.	2

Pagal valstybės institucijų ir įstaigų informacinių sistemų elektroninės informacijos techninius saugos reikalavimus, sistemos turi atitikti tokiems kompleksiškumo reikalavimams: ketvirtosios kategorijos informacinės sistemos turi būti sudarytos iš dviejų ir daugiau modulių, trečiosios kategorijos – iš trijų ir daugiau modulių, antrosios kategorijos – iš penkių ir daugiau modulių, pirmosios kategorijos – iš septynių ir daugiau modulių. Priimama, kad yra modeliuojama trečiosios kategorijos sistema turi penkis modulius m. Kiekvienas modulis atsakingas už tam tikros sistemos dalį, pvz. operacinė sistema, programinė įranga, pašto serveris, ugniasienė, maršrutizatorius ir t. t.

3.2. Informacinės sistemos modelio charakteristikos

Šiame tyrime priimama, kad grėsmės kyla tik iš informacinės sistemos perimetro išorės. Žemiau aprašytam tyrimui buvo naudotos trečiosios kategorijos informacinės sistemos. Taip pat yra aptartos galimos grėsmes ir apžvelgti panaudoti saugumo mechanizmai. Panaudotų saugumo mechanizmų skaičius kiekvienam informacinės sistemos moduliui yra skirtingas.

Saugumo analizė parodė, kad modulio sukompromitavimo aptikimo intervalas yra Δt_d, saugumo modulių svarbą parodo $w(m)$, o skirtingų saugumo modulių panaudojimo dažnis parodo panaudojimo tikimybes $P_{\mathrm{M}}(m)$. Incidentų tikimybės yra pasiskirsčiusios pagal atitinkamas tikimybes: konfidencialumas ($P_{\mathrm{C}m}(j)$), vientisumas ($P_{\mathrm{I}m}(j)$) ir pasiekiamumas ($P_{\mathrm{A}m}(j)$) atitinkamai skirtingiems

moduliams ir grėsmių sunkumams $P_m(j)$. Modeliuojamos informacinės sistemos charakteristikos, apibrėžtos pagal rizikos analizę, yra pateiktos 3.2 lentelėje.

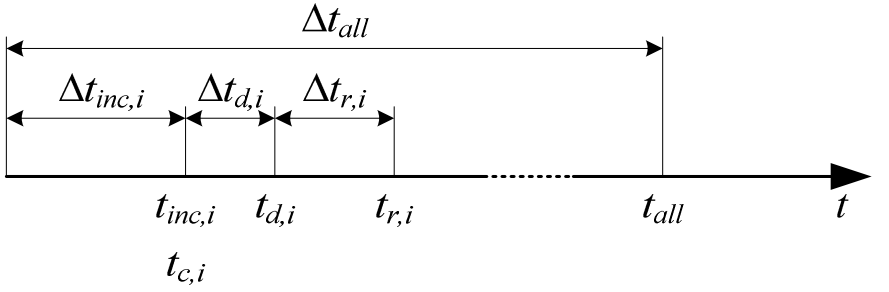
3.2 lentelė. Modeliuojamos informacinės sistemos charakteristikos
Table 3.2. Modelled information system characteristics

m	1			2			3			4			5		
Δt_d	0,1			0,5			1			2			4		
w(m)	0,5			0,3			0,1			0,05			0,05		
j	$P_m(j)$														
	K	V	P	K	V	P	K	V	P	K	V	P	K	V	P
3	,04	,04	,05	,04	,04	,05	,04	,04	,05	,18	,15	,10	,18	,15	,10
2	,05	,05	,06	,05	,05	,06	,05	,05	,06	,23	,20	,12	,23	,20	,12
1	,06	,06	,08	,06	,06	,08	,06	,06	,08	,29	,25	,15	,29	,25	,15

čia ,04 = 0,04

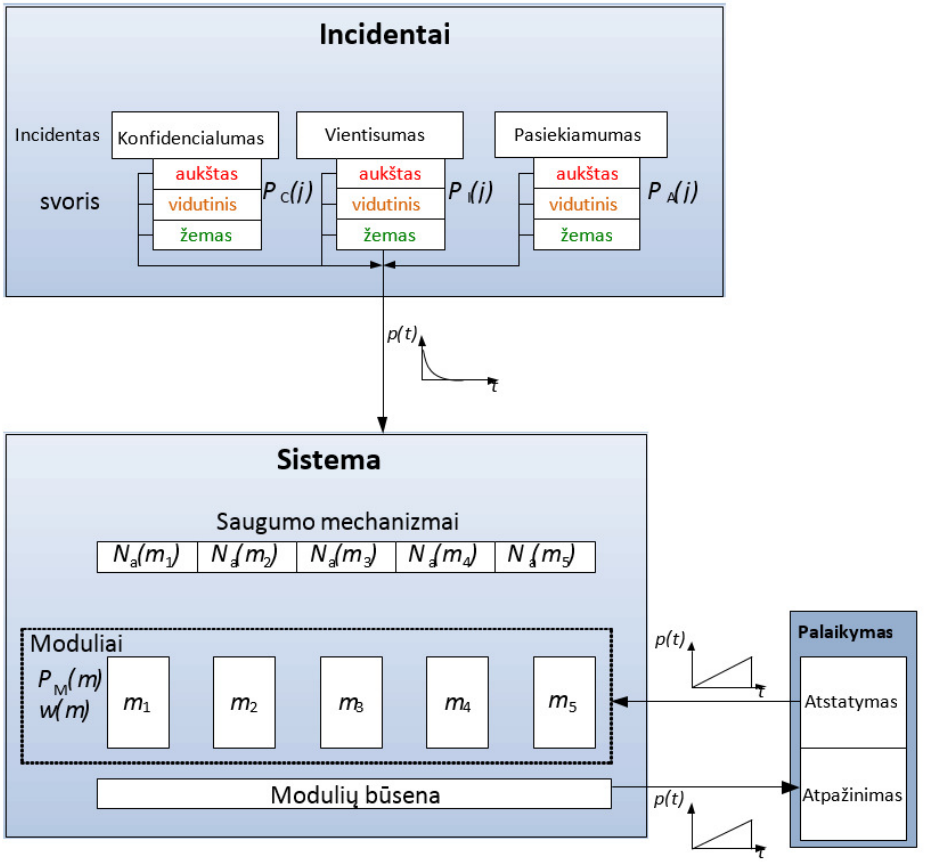
3.3. Modelio sudarymas

Informacinė sistema – tai paskirstytas kompiuterių tinklas su apibrėžtomis ribomis, į kurias yra nukreipiami incidentai *i,* kurie atlaikomi laiko intervale $\Delta t_{inc,i}$, o laiko momentu $t_{c,i}$ vienas ar daugiau modulių yra sukompromituojami. Sistemos pažeidimas yra aptinkamas po laiko $\Delta t_{d,i}$, tada informacinės sistemos būsena yra atstatoma po laiko intervalo $\Delta t_{r,i}$. Informacinė sistema yra modeliuojama laiko intervale Δt_{all}, kuris yra pakankamai ilgas, kad galėtų pasirodyti visi įmanomi įvykiai (3.1 pav.).



3.1 pav. Įvykių informacinėje sistemoje pasirodymas
Fig. 3.1. Information system security events

Įvykių pasirodymas modeliuojamoje sistemoje yra stochastinis procesas, tai ne genetinis algoritmas, kuris dažnai yra naudojamas modeliuojant atakas į sistemas. Incidentai yra sugrupuoti pagal grėsmes ir gali turėti skirtingus sunkumo lygius j, pirmasis lygis yra sunkiausias ($j = 1$). Incidentai yra pasiskirstę pagal Puasono dėsnį (Sanders *et al.* 2010, Moore *et al.* 2001) ir yra nukreipti į tam tikrą modeliuojamos sistemos modulį m, atsižvelgiant į jo panaudojimo dažnį. Visi sistemos moduliai yra apsaugoti saugumo mechanizmais N_a, kurie nustatyti iš saugumo rizikos analizės. Informacinės sistemos moduliai turi skirtingą svarbumą, kuris yra išreikštas kaip modulio svarba $w(m)$. Informacinės sistemos modelis yra parodytas 3.2 paveiksle. Modelio parametrai yra parinkti remiantis saugumo rizikos analize arba nustatyti saugumą reguliuojančio įstatymo. Atitinkamos modelio dalys yra atvaizduotos diagramoje (3.2 pav.).

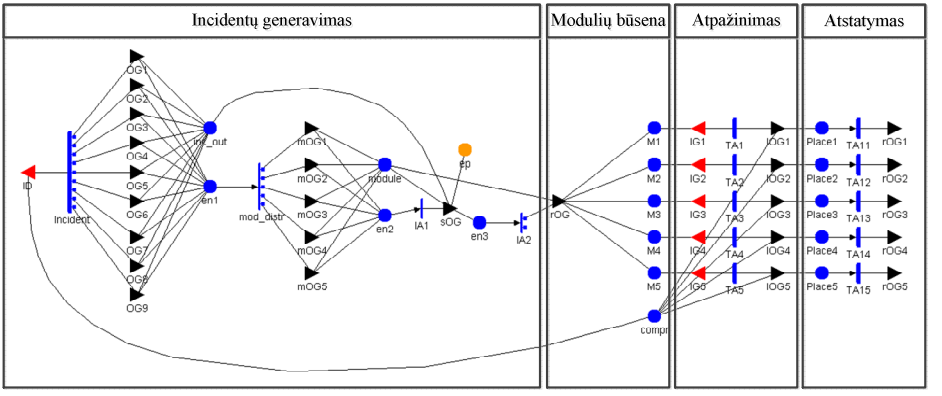


3.2 pav. Informacinės sistemos modelis
Fig. 3.2. information system model

Nuo to ar informacinės sistemos modulis yra sukompromituotas, ar ne, priklauso visos sistemos būsena. Informacinė sistema gali būti vienoje iš penkių būsenų: normali (b_1) – kai nei vienas modulis nėra sukompromituotas; b_2 – kai vienas modulis yra sukompromituotas; kai vienas arba daugiau modulių yra sukompromituotų, sistemos būsenos pasikeitimas yra aptiktas ir sistema yra atstatymo būsenoje (b_3); būsena b_4 – kai daugiau negu pusė modulių yra sukompromituotų; b_5 – kai visi sistemos moduliai yra pažeisti.

Labiausiai yra tikėtina, kad sistemos pažeidimas bus aptiktas greičiau, o sistemos modulio darbas atstatytas per trumpiausią laiko tarpą, dėl to gedimo aptikimui yra naudojamas trikampio pasiskirstymo dėsnis.

Informacinės sistemos modeliavimui buvo naudojami stochastinės veiklos tinklai (angl. *Stochastic Activity Network*), kurie yra labai panašūs į stochastinius Petri tinklus. Modeliavimui panaudotas *Mobius* įrankis. Informacinės sistemos modelis, gautas panaudojus vidinį *Mobius* grafinį redaktorių, parodytas 3.3 paveikslo a) dalyje.



a)

b)

3.3 pav. Informacinės sistemos išliekamumo modelis, atvaizduotas panaudojus SAN: a) modelis, b) SAN baziniai elementai

Fig. 3.3. Information system survivability simulation model by using SAN: a) the model, b) SAN primitives

Stochastinės veiklos tinklai susideda iš tokių bazinių elementų, kaip aikštelė, veikla, įėjimo vartai ir išėjimo vartai (3.3 pav., b) (Sanders *et al.* 2010). Aikštelės savyje laiko žetonus, kurie atvaizduoja vertę arba būseną. Laikinės veiklos gali perkelti žetonus iš vienos aikštelės į kitą pagal tam tikrą laiko pasiskirstymo funkciją. Momentinės veiklos perkelia žetoną akimirksniu. Atvejai (pažymėti taškais prie veiklų) yra naudojami tikimybių atvejų apibrėžimui, įėjimo vartai – predikatų aktyvavimui funkcijų įvykdymui, išėjimo vartai naudojami funkcijų aprašymui. Laikas tarp aktyvuotos ir išpildytos veiklos gali būti skirtingas ir yra priklausomas nuo pasiskirstymo dėsnio, o pasiskirstymo parametrai gali būti būsenų funkcijos.

SAN modelį sudaro keturios pagrindinės dalys: incidentų generavimas, modulių būsenos, incidentų atpažinimas ir sistemos atstatymas. Laikinė veikla „Incident“ yra naudojama incidentų generavimui. Buvo padaryta prielaida, kad incidentai yra nepriklausomi ir jų pasirodymas yra pasiskirstęs pagal Puasono dėsnį. Taip pat remiantis saugumo rizikos analize buvo priimta prielaida, kad sistema yra atakuojama tris kartus per parą. Laikinės veiklos „Incident“ atvejai atvaizduoja incidentų sunkumą skirtingoms grėsmėms (konfidencialumui, vientisumui ir pasiekiamumui), kurios apskaičiuotos iš saugumo rizikos analizės. Priklausomai nuo to koks incidento tipas ir jo sunkumas pasireiškia, veikla „Incident“ įvykdoma, vieni iš išėjimo vartų OG1–OG9 įrašo į aikštelę „inc_out“ numerį, kuriame yra užkoduotas incidento tipas ir jo sunkumo lygis. Momentinė veikla „mod_distr“ yra įvykdoma, kai aikštelėje en1 yra įrašomas „1“. Ši veikla yra naudojama aprašyti informacinės sistemos modulio panaudojimo tikimybę $P_M(m)$. Priklausomai nuo tikimybės $P_M(m)$ išėjimo vartai mOG1–mOG5 pakeičia „module“ aikštelės reikšmę, kur įrašomas įtakojamo modulio numeris. Momentinė veikla IA1 yra įvykdoma, kai aikštelėje en2 yra įrašomas „1“. Išėjimo vartai sOG priklausomai nuo sugeneruoto incidento tipo, jo sunkumo lygio ir įtakoto modulio, įrašo atitinkamą reikšmę į ep aikštelę. Ši reikšmė atvaizduoja modulio sukompromitavimo tikimybę ir yra naudojama tolimesniam pasiskirstymui IA2 – veiklai, kuri turi du atvejus. Jeigu yra pasirenkamas pirmas atvejis, tada išėjimo vartai rOG priklausomai nuo modulio įrašo „1“ į atitinkamą aikštelę (M1–M2), kuri atvaizduoja, kad modulis yra sukompromituotas, kitaip incidentas yra nesėkmingas. Įėjimo vartai IG1–IG5 yra naudojami nustatant laikinės veiklos TA1–TA5 įvykdymo predikatus. Laikinės veiklos TA1–TA5 atvaizduoja incidento aptikimą, TA11–TA15 atvaizduoja sistemos modulio atstatymą ir naudoja tikimybės trikampio pasiskirstymą.

Informacinės sistemos išliekamumas S yra universali skaitinė charakteristika parodanti sistemos galimybę atlikti jai pateiktas funkcijas tam tikroje aplinkoje, kur yra įtaka teikiamoms paslaugoms (Moore *et al.* 2001, Moitra *et al.* 2000).

Sistemos išliekamumas būsenoje b_1 per visą laiko periodą Δ_{tall} yra vadinamas maksimaliu išliekamumu S_{max}:

$$S_{\max} = \frac{\Delta t_{b1}}{\Delta t_{all}} . \tag{3.1}$$

Jeigu sistema išlieka būsenoje b_4, kai pusė jos modulių yra veikiančių per visą laiko periodą Δt_{all}, tokia charakteristika yra vadinama vidutine išliekamumo charakteristika S_{mid}:

$$S_{mid} = \frac{\Delta t_{b4}}{\Delta t_{all}} . \tag{3.2}$$

Turi būti įvertinta skirtinga funkcijų arba modulių svarba sistemos darbui. Sistemos išliekamumas S gali būti aprašytas formule:

$$S = \sum_m w(m)S(m),\ 0 \le S(m) \le 1,\ \sum w(m) = 1,\ 0 \le w(m) \le 1, \tag{3.3}$$

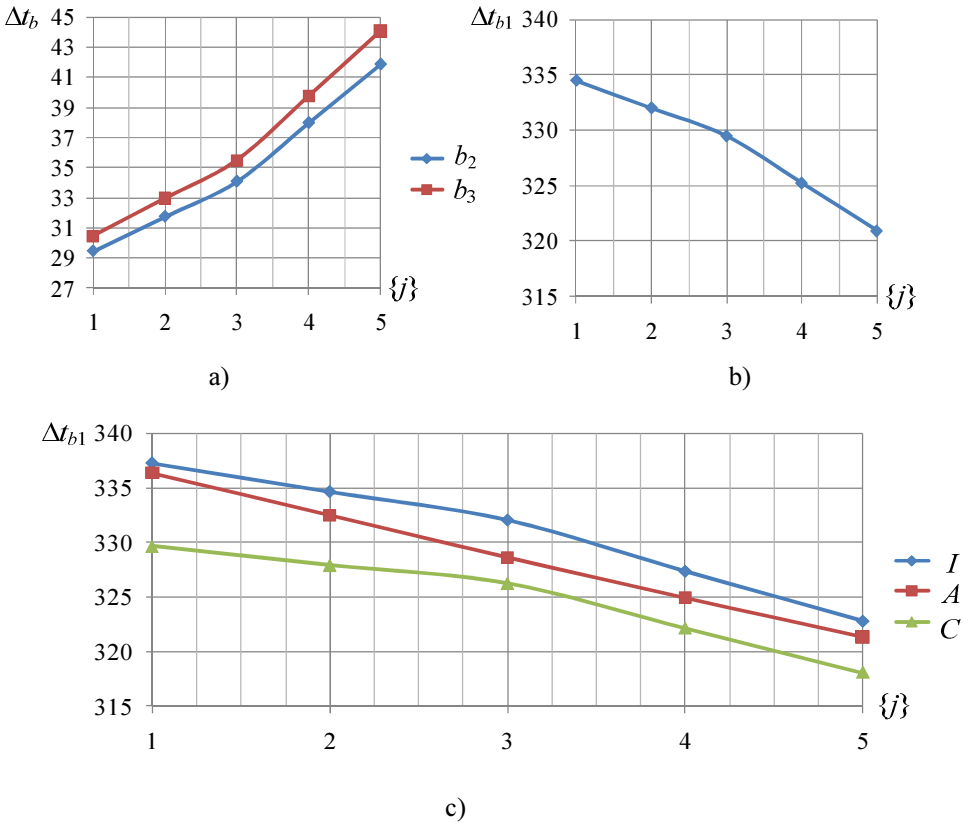
čia $S(m)$ informacinės sistemos modulio m išliekamumas, o $w(m)$ modulio svarba.

Incidentų sunkumo įtaka modeliuotai sistemai įvertinama panaudojant skirtingus incidentų rinkinius J – nuo vidutinio sunkumo $j = 1$ iki pačio didžiausio sunkumo $j = 3$ (3.2 lentelė).

3.2 lentelė. Incidentų sunkumo rinkinys
Table 3.2. Incident severity sets

J		1	2	3	4	5
j	1	0	0	0	0,5	1
	2	0	0,5	1	0,5	0
	3	1	0,5	0	0	0

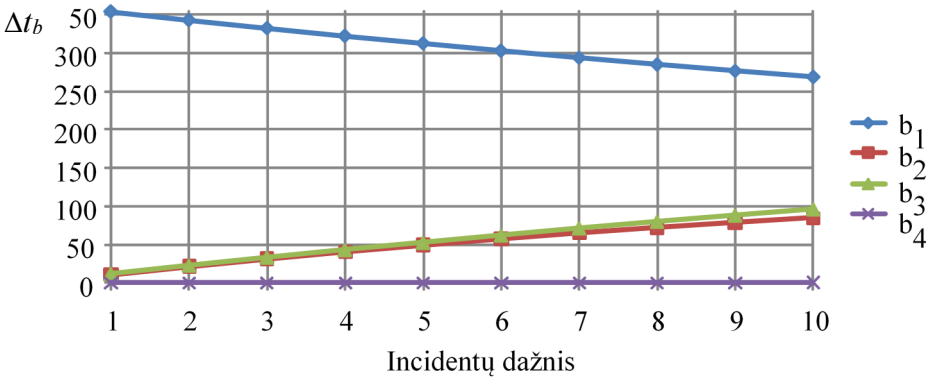
Sistemos laikas sukompromituotoje būsenoje ilgėja didėjant incidentų sunkumui (3.4 pav., a), o sistemos buvimas normalioje būsenoje trumpėja didėjant incidentų sunkumui (3.4 pav., b). Incidentų įtaka skirtingoms grėsmėms yra skirtinga dėl informacinės sistemos modulių apsaugos mechanizmų savybių, kurios priklauso nuo kompiuterių, tačiau sunkaus lygio incidentų įtaka yra panaši (3.4 pav., c).



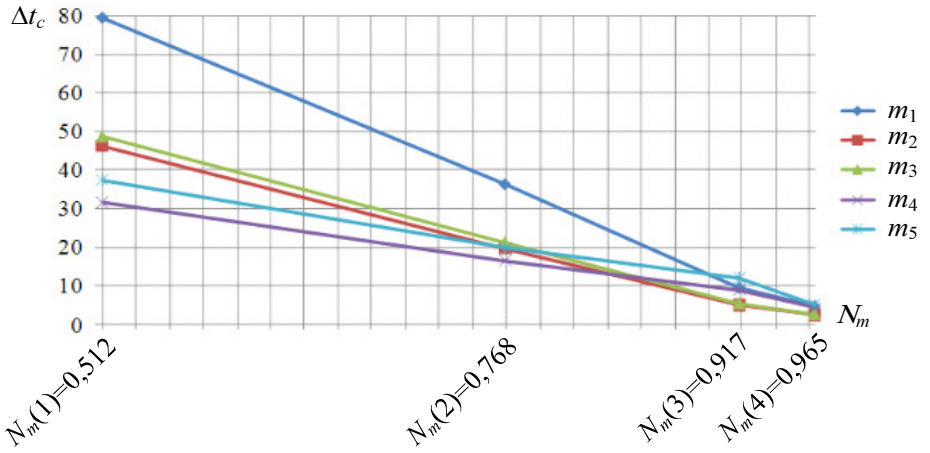
3.4 pav. Incidentų sunkumo įtaka sistemos būsenai: a) b_2 ir b_3; b) b_1; c) b_1 priklausomai nuo grėsmės

Fig. 3.4. Incident Severity Influence on System States: a) b_2 and b_3; b) b_1; c) b_1 according to the threat

Didėjant incidentų pasirodymo dažniui modeliuojamos sistemos tikimybė būti sukompromituotai didėja (3.5 pav., a). Skirtingi informacinės sistemos moduliai yra apsaugoti skirtingais saugumo mechanizmais ir priklausomai nuo sistemos modulio svorio tikimybė būti sukompromituotam yra skirtinga. Kuo saugumą užtikrinantys mechanizmai yra stipresni, tuo tikimybė sistemai likti normalioje būsenoje yra didesnė (3.5 pav., b).



a)

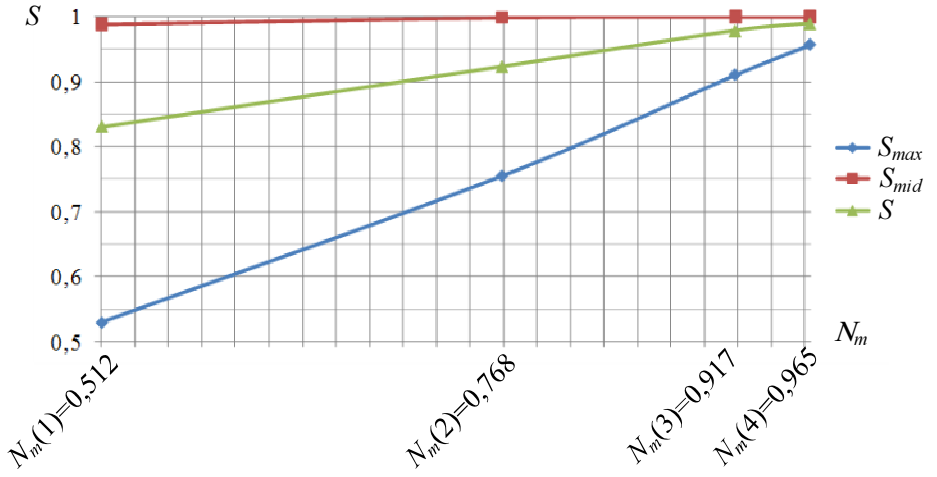


b)

3.5 pav. a) incidentų pasirodymo dažnio įtaka sistemos būsenai; b) saugumo mechanizmų rinkinių įtaka normaliai sistemos būsenai pagal sistemos modulius
Fig. 3.5. a) Incident Occurrence Interval Influence on System States; b) Protection Mechanism Set Influence on Normal System State according to the Module

Išliekamumas yra kiekybinė informacinių sistemų saugumo charakteristika, kuri yra pavaizduota 3.6 paveikslo diagramoje. Maksimalus išliekamumas S_{max} – tai tikimybė, kad informacinė sistema po incidento liks normalioje būsenoje. S_{mid} – tai tikimybė, kad pusė informacinės sistemos modulių liks normalioje būsenoje. Informacinės sistemos išliekamumas S parodo vidutinę išliekamumo reikšmę, kuri geriausiai atvaizduoja informacinės sistemos saugumo mechanizmų įtaką modeliuojamai sistemai.

Skirtingi saugumo mechanizmų rinkiniai $N_{m(n)}$, kur n yra saugumo mechanizmų kiekis, skiriasi tarpusavyje, o jų reikšmė yra išreikšta saugumo mechanizmų su trečiosios kategorijos informacinės sistemos visų saugumo mechanizmų skaičiumi.



3.6 pav. Saugumo mechanizmų įtaka sistemos išliekamumui
Fig. 3.6. Survivability characteristics

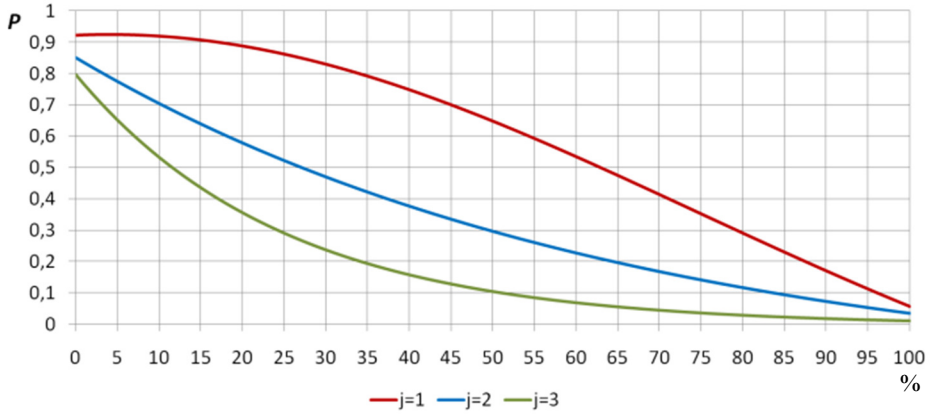
Stochastinės veiklos tinklų modelyje buvo panaudota viena informacinės sistemos kategorija, o modelio parametrai buvo nustatyti iš rizikos analizės rezultatų. Toliau buvo sukurti keturi informacinės sistemos modeliai (po vieną modelį sistemos kategorijai). Šiame poskyryje pateikiamas būdas, leidžiantis apskaičiuoti sistemos sukompromitavimo tikimybę, taikomas informacinės sistemos išliekamumo modeliavimui. Taip pat yra atliktas keturių informacinės sistemos kategorijų palyginimas. Kiekviena iš informacinės sistemos kategorijų skiriasi modulių skaičiumi: pirmoji kategorija turi 7 modulius, antroji – 5 modulius, trečioji – 3 modulius ir ketvirtoji – 2 modulius.

3.4. Sistemos sukompromitavimo tikimybė

Sistemos išliekamumo modeliavimui būtina žinoti sistemos sukompromitavimo tikimybę, pvz. tikimybę, kad tam tikras incidentas sukompromituos sistemą. Ši tikimybė priklauso nuo incidento sunkumo ir į sistema įdiegtų apsaugos mechanizmų kiekio. Geriausias būdas yra sukaupti statistiką apie incidentus, įtakojančius sistemą per tam tikrą laiko periodą, bet dažniausiai tokia statistinė

informacija yra neprieinama, pvz. sistema yra vystymo stadijoje arba ką tik įdiegta. Šiame skyriuje sistemos sukompromitavimo tikimybėms gauti yra siūloma naudoti teorines charakteristikas.

Grafikai pavaizduoti 3.7 paveiksle parodo kaip sistemos sukompromitavimo tikimybė priklauso nuo įdiegtų apsaugos mechanizmų skaičiaus. Čia parodyti trys grafikai skirti kiekvienam incidento sunkumo lygiui. Pirmas sunkumo lygis yra labiausiai įtakojantis sistemą (sunkiausias), tai reiškia, kad kuo sunkesnis incidentas, tuo sistemos sukompromitavimo tikimybė didesnė.



3.7 pav. Sukompromitavimo tikimybės priklausomybė nuo įdiegtų saugumo mechanizmų skaičiaus

Fig. 3.7. Compromise probability dependency on the amount of implemented security mechanisms

Sukompromitavimo tikimybių kreivės buvo pasirinktos remiantis šiais teiginiais:

- bazinių saugumo mechanizmų kiekis turi būti didesnis tam, kad kompensuoti sunkiausių incidentų įtaką;
- jeigu incidentas yra lengvesnis, tai mažesnis saugumo mechanizmų kiekis yra reikalingas tam, kad kompensuoti incidento įtaką;
- kai saugumo mechanizmų kiekis yra mažas arba jų nėra, nepriklausomai nuo incidentų sunkumo, jų įtaka panaši ir sistema bus sukompromituota.

Naudojant visus įmanomus saugumo mechanizmus visų sunkumų incidentų įtaka yra maždaug tokia pat. Žemiau pateikiamos formulės, kurios aprašo 3.7 paveikslo kreives (kur $\beta_{Th} = 0$ ir $\alpha_{mi} = 0$):

$$P(x)_{j=1} = \left(\frac{25}{7\sqrt{2\pi}} \cdot e^{-\frac{\left(x-4+\alpha_{m_i}\right)^2}{9800}} \right) - 0{,}5 + \beta_{Th}, \qquad (3.4)$$

$$P(x)_{j=2} = \left(1{,}05 \cdot e^{-0{,}015\cdot\left(x+\alpha_{m_i}\right)}\right) - 0{,}2 + \beta_{Th}, \quad (3.5)$$

$$P(x)_{j=3} = \left(0{,}8 \cdot e^{-0{,}04\cdot\left(x+\alpha_{m_i}\right)}\right) - 0{,}004 + \beta_{Th}, \quad (3.6)$$

čia x – įdiegtų saugumo mechanizmų skaičius [0–100]; β_{Th} – koeficientai grėsmės tipo įvertinimui (β_C – konfidencialumas, β_A – pasiekiamumas, β_I – vientisumas); α_{mi} – koeficientai įvertinantys modulio atsparumą incidentams.

Šios formules yra naudojamos modelyje tam, kad nustatyti tikras sukompromitavimo tikimybių reikšmes. Grėsmės tipas yra įvertinamas panaudojus β_{Th} koeficientą. Šie koeficientai parodo saugumo mechanizmų savybes prieš tam tikrus grėsmių tipus ir gali būti teigiami arba neigiami, pvz. jeigu tam tikro saugumo mechanizmo savybės yra tinkamesnės apsaugai nuo grėsmes konfidencialumui negu vientisumui, tada β_C koeficientas turės neigiamą reikšmę (žemesnę sukompromitavimo tikimybę), o β_I koeficientas turės teigiamą reikšmę (aukštesnę sukompromitavimo tikimybę).

Sistemos modulių savybės apsisaugojimui nuo grėsmės yra skirtingos priklausomai nuo modulio. Skirtumas yra įvertinamas panaudojus α_{mi} koeficientą, kuris taip pat gali turėti teigiamą arba neigiamą reikšmę. Sistemos modulio svarba sistemai yra parodomas svoriu $w(m)$, kai įgyvendintas saugumo mechanizmas yra kruopščiai pritaikytas, tai reiškia α_{mi} koeficientas šiam konkrečiam moduliui bus teigiamas, t. y. sukompromitavimo tikimybė bus mažesnė. Pasiūlytos formulės leidžia adaptuoti sistemos sukompromitavimo tikimybes remiantis grėsmės tipu, sunkumu ir saugumo mechanizmų rinkiniu. Formulė gali būti lengvai modifikuojama ir pritaikoma realios kompiuterių sistemos poreikiams.

Modeliavimo rezultatai

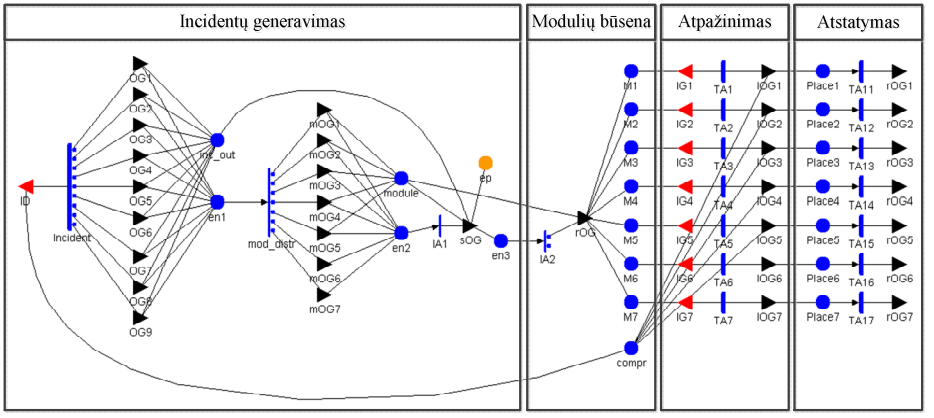
Išliekamumas – tai atsparumo atakoms lygis, kuris parodo sistemos galimybes toliau funkcionuoti po įvykusios atakos, tai kiekybinė informacinės sistemos saugumo charakteristika. Pagrindiniu modeliavimo parametru yra modeliavimo trukmė – 365 dienos, t. y. 1 metai. Informacinės sistema yra atakuojama 3 kartus per dieną. Atakos yra nepriklausomos ir pasiskirsčiusios eksponentiškai.

Atakų pasirodymų tikimybės:

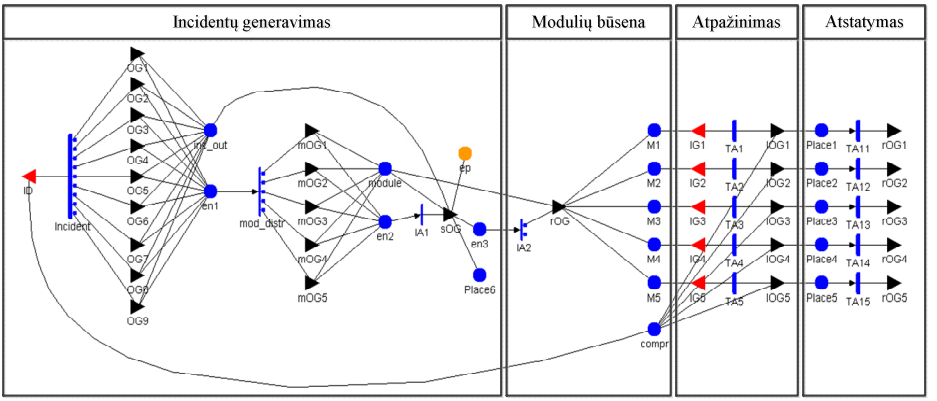
- *konfidencialumas* – $P_C(j = 1) = 0{,}04$, $P_C(j = 2) = 0{,}1$, $P_C(j = 3) = 0{,}2$;
- *vientisumas* – $P_I(j = 1) = 0{,}01$, $P_I(j = 2) = 0{,}05$, $P_I(j = 3) = 0{,}1$;
- *pasiekiamumas* – $P_A(j = 1) = 0{,}1$, $P_A(j = 2) = 0{,}1$, $P_A(j = 3) = 0{,}3$;

čia $\sum P_C(j)+\sum P_I(j)+\sum P_A(j)=1$. Šios reikšmės buvo pasirinktos tik modeliavimo tikslams.

Buvo sudaryti keturių kategorijų sistemų modeliai. Sistemų modeliai pateikiami 3.8 ir 3.9 paveiksluose.



a)



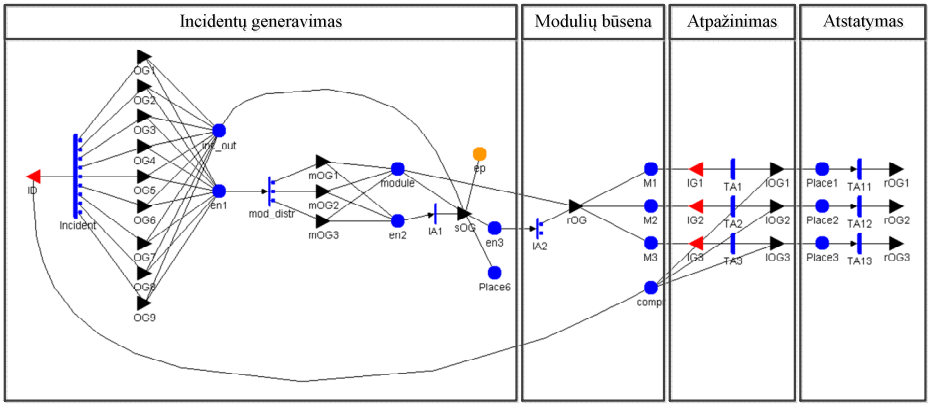
b)

3.8 pav. Informacinių sistemų modeliai: a) 1-osios kategorijos, b) 2-osios kategorijos
Fig. 3.8. Information system survivability simulation model by using SAN: a) the 1st category model, b) the 2nd category model

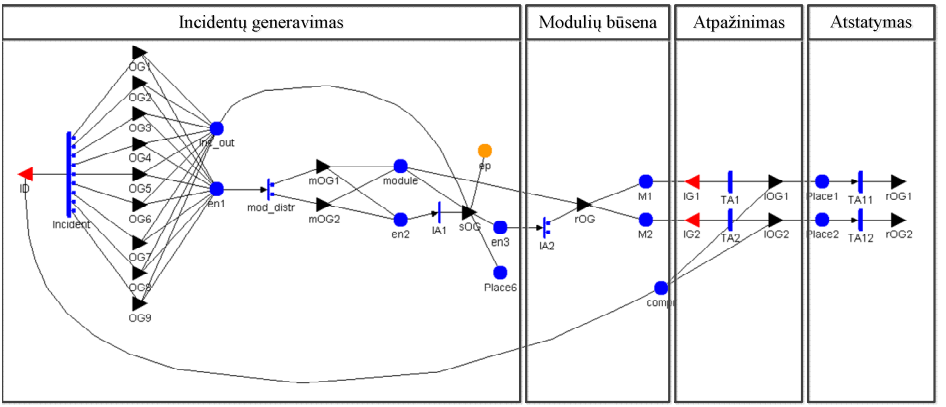
Pirmosios ir antrosios kategorijos sistemos turi didesnius pasiekiamumo reikalavimus, jų išliekamumo charakteristika, gauta modeliavimo metu, pavaizduota 3.10 paveikslo a) dalyje. Šių kategorijų sistemų išliekamumas mažiau priklauso nuo saugumo mechanizmų.

Kaip matyti, išliekamumo charakteristikų skirtumas tarp sistemos, kurioje įdiegti visi įmanomi saugumo mechanizmai ir kurioje jų skaičius mažesnis,

skiriasi nežymiai. Bet šį skirtumą pavertus dienomis, matysime 2,9 dienų skirtumą. Kaip buvo pateikta ankščiau, pasiekiamumo reikalavimai pirmosios kategorijos sistemoms yra 99 %, taigi 2,9 dienų skirtumas yra gan apčiuopiamas. Papildomai mažas skirtumas gali būti paaiškintas atsižvelgiant į modelio charakteristikas.



a)



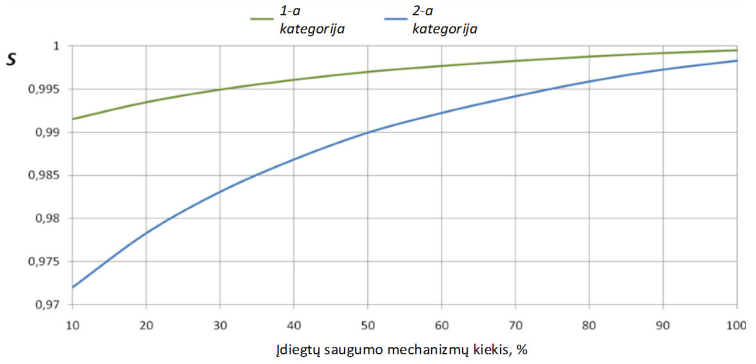
b)

3.9 pav. Informacinių sistemų modeliai a) 3-osios kategorijos, b) 4-osios kategorijos
Fig. 3.9. Information system survivability simulation model by using SAN: a) the 3th category model, b) the 4th category model

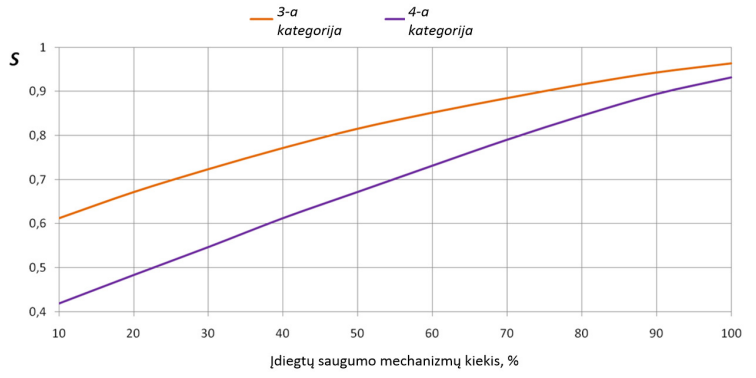
Sistemos atstatymo laikas modeliavimo metu buvo pastovus, reiškia ir esant mažam mechanizmų kiekiui, atstatymo laikas likdavo nepakitęs. Sekantis

modelio patobulinimas būtų atstatymo laiko nuo įdiegtų saugumo mechanizmų kiekio įvertinimas.

Negalima palyginti 1-osios ir 2-osios sistemų kategorijų su 3-čiąja ir 4-tąja kategorijomis, nes jų pasiekiamumo reikalavimai yra skirtingi. Pasiekiamumo reikalavimai 3-osios ir 4-osios kategorijos sistemoms galioja tik darbo dienomis, tuo tarpu 1-osios ir 2-osios kategorijos sistemoms galioja visomis savaitės dienomis. Paveikslo 3.10 dalyje b) parodyta išliekamumo charakteristika 3-osios ir 4-osios kategorijos sistemoms. Iš pateiktų charakteristikų matyti, kad 3-osios ir 4-osios kategorijų sistemų išliekamumo charakteristikos yra daug tiesiškesnės negu 1-osios ir 2-osios kategorijų sistemų. Tai reiškia, kad jų išliekamumas labiau priklauso nuo įdiegtų saugumo mechanizmų kiekio.



a)



b)

3.10 pav. Išliekamumo charakteristikos: 1) 1 ir 2 sistemų kategorijos, b) 3 ir 4 sistemų kategorijos

Fig. 3.10. Survivability characteristics: a) 1st and 2nd system categories, b) 3th and 4th system categories

Pateiktas modelis leidžia įvertinti informacinės sistemos išliekamumą priklausomai nuo grėsmės tipo. Sukurtas modelis yra lankstus, lengvai pakeičiamas ir gali būti pritaikytas bet kokiai informacinei sistemai. Pirmosios ir antrosios kategorijos informacinėms sistemoms yra keliami didesni išliekamumo reikalavimai, ką patvirtino išliekamumo ir modeliavimo rezultatai – išliekamumo charakteristika yra mažiau priklausoma nuo saugumo mechanizmų kiekio. Modeliuojant buvo gauti sistemų išliekamumo charakteristikų skirtumai dėl saugumo mechanizmų kiekio: pirmajai kategorijai 2,9 dienos, antrajai – 9,5 dienos, trečiajai – 128 darbo dienos, ketvirtajai – 187 darbo dienos. Taip pat modeliavimo rezultatai parodė, kad grėsmės atpažinimo trukmė priklauso nuo įdiegtų saugumo mechanizmų kiekio. Tolesniame tyrime būtina įvertinti informacinės sistemos atstatymo laiko priklausomybę nuo įdiegtų saugumo mechanizmų kiekio.

3.5. Trečiojo skyriaus išvados

1. Modeliuojamos informacinės sistemos saugumas labiausiai priklauso nuo incidentų pasirodymo tikimybės, apsaugos mechanizmų stiprumo, tuo tarpu incidentų lygis turi mažiausiai įtakos informacines sistemos saugumui.
2. Pirmosios ir antrosios kategorijų informacinės sistemos yra projektuojamos su didesniais pasiekiamumo reikalavimais ir šių kategorijų, dėl to gautos išliekamumo charakteristikos yra mažiau priklausomos nuo incidentų tipo ir dažnio negu trečiosios ar ketvirtosios kategorijos informacinių sistemų išliekamumo charakteristikos.
3. Sukurtas informacinių sistemų saugumo modelis leido apskaičiuoti sistemos pasiekiamumo skirtumą dienomis tarp atvejų, kai įdiegtų saugumo mechanizmų skaičius yra mažiausias ir didžiausias: pirmosios kategorijos sistema nebūtų pasiekiama – 2,9 dienas, antrosios – 9,5 dienas, trečiosios – 128 darbo dienas, ketvirtosios – 187 darbo dienas.
4. Kuo didesni pasiekiamumo reikalavimai, tuo mažiau sistemos išliekamumas priklauso nuo panaudotų apsaugos mechanizmų.

4

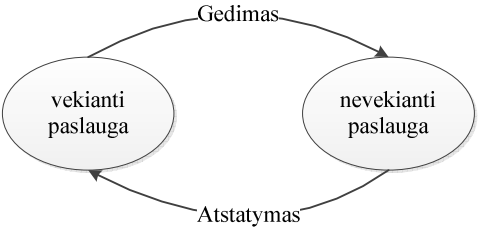
Akademinio tinklo informacinių sistemų saugumo lygio įvertinimas

Šiame skyriuje aptartos informacinių sistemų patikimumo sudedamosios dalys. Atlikta realaus tinklo informacinių sistemų saugumo incidentų ir saugumo reikalavimų analizė, iš kurios gauti statistiniai pasiskirstymai buvo pritaikyti informacinės sistemos saugumo modeliavime. Pasinaudojus gautais statistiniais pasiskirstymo duomenimis gauta akademinio tinklo informacinės sistemos saugumo lygio įvertinimo charakteristika, iš kurios padarytos išvados ir sistemos saugumo lygio tyrimas. Šio skyriaus medžiaga buvo pristatyta Jaunųjų mokslininkų konferencijoje 2014 metais.

4.1. Informacinių sistemų patikimumas ir jo sudedamosios dalys

Skirtingose sistemose turi būti užtikrintos skirtingos savybės, pvz. vidutinis kliento aptarnavimo laikas, rezultato patikimumas, gebėjimas išvengti sutrikimų, kas galėtų katastrofiškai paveikti sistemos darbą. Šioms charakteristikoms apibūdinti buvo įvesta bendra sąvoka – patikimumas (angl. *Dependability*), kuri apjungia visą eilę reikalavimų į vieningą susistemintą struktūrą.

Patikimumas – tai sistemos gebėjimas teikti nustatytas paslaugas. Paslaugų teikimas gali būti normalus arba sutrikdytas. Sistemos sutrikimas – tai perėjimas iš normalaus paslaugos teikimo režimo į sutrikdyto paslaugos teikimo režimą. Sistemos atstatymas – tai grįžimas į normalią paslaugos teikimo būseną (Avižienis 2001). Perėjimas tarp būsenų parodytas diagramoje (4.1 pav.).



4.1 pav. Būsenų perėjimo diagrama
Fig. 4.1. Diagram of state transition

Sistemų patikimumas susideda iš trijų pagrindinių dalių:

- veiksniai (angl. *Treats*) – tai veiksniai įtakojantys sistemų patikimumą;
- matas (angl. *Attributes*) – sistemų patikimumo įvertinimo būdas;
- priemonės (angl. *Means*) – veiksniai, leidžiantys padidinti sistemų patikimumą.

Iš veiksnių veikiančių sistemos patikimumą galima išskirti tris grupes:

- defektas (angl. *Fault*) – tai defektas sistemoje, kurio buvimas ne visada yra gedimo priežastimi, pvz. sistemoje yra defektas, bet jeigu sistemos panaudojimas nepriveda prie funkcijos su defektu įvykdymo, tai šio defekto buvimas nesukels sistemos gedimo;
- klaida (angl. *Error*) – tai yra prieštaringas sistemos veikimas tarp būsimo sistemos elgesio ir jos faktinio elgesio. Klaidos atsiranda sistemos veikimo metu, kai sistemoje atsiranda nenumatyta būsena, kurios priežastimi galėjo būti defektas. Klaidos dažnai ieškomos programos derinimo metu;
- gedimas (angl. *Failure*) – tai tokia sistemos būsena, kai jos veikimas prieštarauja specifikacijai. Klaida ne visada gali tapti gedimo priežastimi, jeigu yra naudojamos gedimo prevencinės priemonės.

Gedimas gali pasireikšti sistemos ribose. Visus veiksnius sieja grandinė defektas – klaida – gedimas, defektas sukelia klaidą, klaida sukelia kitą klaidą arba gedimą. Jeigu klaida išeina iš sistemos ribų, tai sukelia sistemos arba jos modulio gedimą (Avižienis 2001).

Iš sistemų saugumo pusės veiksniai galintys paveikti sistemą galėtų būti: fizinis įrangos gedimas, defektai atsiradę sistemos kūrimo metu, vartotojo

padarytos klaidos, sistemos susidėvėjimas, gamtos katastrofos, programinės klaidos, piktavališka veikla, virusai, kirminai, trojos arkliai, atsisakymo aptarnauti atakos ir daug kitų.

Matai, kitaip vadinami atributai, yra sistemos kokybiniai rodikliai. Jie gali būti naudojami patikimumo įvertinimui naudojant kokybinius arba kiekybinius matavimo vienetus. Pagal (Avižienis 2001) šaltinį, patikimumą galime įvertinti tokiais matais:

- pasiekiamumas (angl. *Avialability*) – tai sistemos pasiruošimas teikti servisą;
- patikimumas (angl. *Reliability*) – tai sistemos nepertraukiamo serviso teikimas;
- sauga (angl. *Safety*) – tai kritinių padarinių stoka vartotojui arba sistemai;
- konfidencialumas (angl. *Confidentiality*) – tai neteisėto informacijos atskleidimo nebuvimas;
- vientisumas (angl. *Integrity*) – tai sistemos pokyčio nebuvimas;
- prižiūrimumas (angl. *Maintanability*) – tai sistemos gebėjimas būti valdomai.

Iš visų matavimo vienetų pasiekiamumas ir patikimumas leidžia kiekybiškai nustatyti sistemos būseną tiesiogiai atliekant matavimus, tuo tarpu kiti matavimo vienetai yra subjektyvūs. Pavyzdžiui sauga negali būti išmatuota tiesiogiai, nes tai yra subjektyvus įvertinimas, kuris reikalauja papildomų svarstymų įvertinant pasitikėjimo lygį, tuo tarpu patikimumas gali būti išmatuotas kaip gedimų skaičius per periodą. Saugumas (angl. *Security*) susideda iš pasiekiamumo, patikimumo, saugos, vientisumo ir konfidencialumo.

Kaip jau buvo minėta, egzistuoja defektas – klaida – gedimas grandinė. Priemonės leidžiančios pertraukti šią grandinę ir padidinti sistemos patikimumą yra atskira patikimumo sudedamoji dalis (Avižienis 2004). Šios priemonės suskirstytos į keturias dalis:

- gedimo prevencija (angl. *Fault prevention*) – numato, kaip apsisaugoti nuo galimo sistemos gedimo;
- gedimo tolerancija (angl. *Fault tolerance*) – tai sistemos patikimumo užtikrinimas žinant esamus defektus;
- gedimo šalinimas (angl. *Fault removal*) – numato pašalinti potencialaus gedimo šaltinį arba jo įtaką sistemai;
- gedimo nuspėjimas (angl. *Fault forecasting*) – tai gedimo dabartinėje sistemos būsenoje ir ateityje prognozė.

Iš dalies visas šias priemones apima sistemos projektavimo etapas – tai sistemos architektūros parinkimas ir sistemos patikimumo reikalavimų įgyvendinimo galimybės. Gedimų šalinimo priemonė gali būti įvertinta kaip projektavimo etape taip ir sistemos naudojimo metu. Sistemos architektūros

parinkimo metu galėtų būti numatytos sekančios saugumo didinimo priemonės: aparatinės ir programinės, fizinis įrangos dubliavimas, informacijos atsarginės kopijos, ugniasienės, atakų atpažinimo sistemos, taip pat sistemos atstatymo planas gedimo atveju. Didelį įnašą sistemos saugumo didinime lemia pastovus sistemos būsenos stebėjimas, tai yra defektų radimas ir šalinimas, sistemos patikrinimai, diagnostika.

4.2. Saugumo incidentai tinkle

Saugumo incidentas – realų ar potencialiai nepageidaujamą poveikį informacinei sistemai ar kompiuterių tinklo veiklai turintis įvykis, kurio rezultatas – apgaulė, nuostoliai ar piktnaudžiavimas, grėsmė informacijai, informacijos nuosavybės praradimas ar žala jai (CERT-LT). Pavyzdžiui, skverbimasis į informacines sistemas, techninių pažeidžiamumų išnaudojimas, kompiuterinių virusų ar kitokios nepageidaujamos programinės įrangos įdiegimas.

Trys pagrindinės dalys, svarbios informacijos saugumui kompiuteriniuose tinkluose yra: konfidencialumas (angl. *Confidentiality*), vientisumas (angl. *Integrity*) ir pasiekiamumas (angl. *Availability*). Sąvokos susijusios su žmonėmis, kurie naudojasi informacijos ištekliais, tai tapatybės patikrinimas, įgaliojimų patikrinimas ir atsakomybės pripažinimas.

Kai informacija yra nuskaitoma arba kopijuojama asmens, neturinčio tam įgaliojimų, rezultatas yra vadinamas konfidencialumo praradimu. Vientisumo praradimas – tai perduodamos tinklu informacijos pakeitimas nenumatytu būdu. Kai informacija ištrinama ir/arba tampa nepasiekiama, tai vadinama informacijos pasiekiamumo praradimu. Atvejis, kai teisėtas vartotojas negali pasiekti tinklo arba atskirų jo paslaugų, jis patiria atsisakymą suteikti paslaugą (angl. *Denial of service*).

Išanalizavus kiekvieno incidento tipą galima sudaryti lentelę, kurioje bus aprašoma kiekvieno incidento ataka sistemos saugumo sudedamosioms dalims. Lentelėje 4.1 pateikiama informacija apie incidentų įtaką konfidencialumui, vientisumui ir pasiekiamumui.

4.1 lentelė. Saugumo incidentų įtaka konfidencialumui, vientisumui ir pasiekiamumui
Table 4.1. Security incident influence on confidentiality, integrity and availability

Įtaka sistemos saugumui			
Saugumo incidento tipas	Konfidencialumas	Vientisumas	Pasiekiamumas
Elektroninės paslaugos trikdymo ataka (angl. *DoS*)		+	+

4.1 lentelės pabaiga

Saugumo incidento tipas	Konfidencialumas	Vientisumas	Pasiekiamumas
Kenkėjiškas programinis kodas (angl. *Malware*)	+	+	+
Trojos arkliai (angl. *Trojan Horse*)	+	+	
Nepageidaujamas elektroninis paštas (angl. *Spam*)		+	+
Prievadų žvalga (angl. *Port Scaning*)	+	+	
Elektroninių duomenų klastojimas (angl. *Phishing*)	+		
Neleidžiamasis prisijungimas (angl. *System Compromise / Intrusion*)	+	+	+
Manipuliacija elektroniniais duomenimis (angl. *Spyware*)	+		+
Socialinė inžinerija (angl. *Social Engineering*)	+	+	+

Tam, kad gauti tinkle esančių saugumo incidentų įtakos konfidencialumui, vientisumui ir pasiekiamumui pasiskirstymą yra reikalinga kompiuterių tinklo saugumą užtikrinanti įranga, kuri fiksuoja saugumo incidentų kiekį ir nustato jų tipą. Geriausiai tam tinkančios sistemos yra atakų atpažinimo sistemos (angl. *Intrusion detection* systems). Tyrimui atlikti buvo pasinaudota atakų atpažinimo sistemos surinktais duomenimis.

Iš atakų atpažinimo sistemos buvo gauti 2012 metų visų 12 mėnesių užfiksuotų incidentų duomenys. Ataskaita apima informaciją apie tokius incidentus, kaip: atpažinti virusai (P1 lentelė, A priedas), nepageidaujamo elektroninio pašto (toliau – *Spam*) ataka P2 lentelė, A priedas), ir atsisakymo aptarnauti atakos (P3 lentelė, A priedas). Ataskaitoje pateikti populiariausių incidentų pavadinimai ir jų pasirodymo skaičius.

Pagal viruso pavadinimą buvo ieškomas detalus viruso aprašymas ir pagal tai buvo nustatoma konkretaus viruso įtaka sistemos konfidencialumui, vientisumui, pasiekiamumui ir atakos lygis. Esant neišskirtam viruso tipui buvo

priimta, kad virusai konfidencialumą, vientisumą ir pasiekiamumą įtakoja vienodai, o atakos lygis taip pat pasiskirstęs po lygiai.

Nepriklausomai nuo *Spam* laiško siuntėjo ir turinio, šio tipo atakos buvo priskiriamos prie atakų, kurios įtakoja sistemos vientisumą ir pasiekiamumą, o atakos lygis priskiriamas prie vidutinio.

Pagal atsisakymo aptarnauti atakų tipą buvo ieškomas atakos detalus aprašymas ir pagal tai buvo nustatoma konkreti įtaka sistemos konfidencialumui, vientisumui, pasiekiamumui ir atakos lygis. Esant neišskirtam atakos tipui buvo priimta, kad visos sistemos saugumo dalys (konfidencialumas, vientisumas ir pasiekiamumas) yra įtakojamos po lygiai, o atakos lygis taip pat padalintas po lygiai.

Atakų atpažinimo sistemoje užtvindymo atakų paketai yra skaičiuojami kaip atskiri įvykiai. Dėl didelio įvykių skaičiaus virusų incidentai sudarytų tik 0,005 %, o *Spam* 0,185 % visų incidentų. Dėl to buvo priimtos tam tikros sąlygos užtvindymo ir *Spam* incidentų normalizavimui. Kad įvykių skaičius atvaizduotų vienos vientisos atakos skaičių, buvo priimta sąlyga, kad užtvindymo, sync / sesijų ir skanavimo atakos atitinkamai sudaro po 10 000, 1 000 ir 100 paketų. *Spam* incidentų skaičius buvo padalintas iš 10. Tokios sąlygos buvo paimtos remiantis realiais užtvindynimo atakų srautais. Dėl *Spam* laiškų buvo priimta, kad kas 100-tasis siunčiamas laiškas pasieks adresato pašto dėžutę. Po tokios sąlygos įgyvendinimo buvo gauta, kad virusai sudaro 29,8 %, *Spam* laiškai – 11,0 %, o užtvindymo atakos – 59,2 % visų atakų.

Atlikus saugumo incidentų analizę, buvo gautos incidentų pasiskirstymo tikimybės pagal incidentų sunkumą: nuo 3 (lengviausi) iki 1 (sunkiausi), ir pagal incidento įtaką konfidencialumui, vientisumui ir pasiekiamumui. Incidentų pasiskirstymo reikšmės parodytos 4.2 lentelėje.

4.2 lentelė. Incidentų pasiskirstymas
Table 4.2. Incident distribution

Grėsmė	Incidento sunkumas		
	1	2	3
konfidencialumui	0,148	0,052	0,061
vientisumui	0,120	0,084	0,132
pasiekiamumui	0,293	0,054	0,055

Tiriamoji informacinė sistema yra sudaryta iš penkių modulių, tokių kaip: m_1 – operacinė sistema, m_2 – programinė įranga, m_3 – elektroninio pašto serveris, m_4 – maršrutizatorius, m_5 – ugniasienė. Modulių svarbą parodo modulio svoris $w(m)$, o skirtingų saugumo modulių panaudojimo dažnį parodo panaudojimo

tikimybė $P_{M(m)}$. Incidentų tikimybės yra pasiskirsčiusios pagal atitinkamas tikimybes: konfidencialumas ($P_{Cm(j)}$), vientisumas ($P_{Im(j)}$) ir pasiekiamumas ($P_{Am(j)}$) atitinkamai skirtingiems moduliams ir grėsmių sunkumams $P_{m(j)}$. Modeliuojamos informacinės sistemos charakteristikos, apibrėžtos pagal rizikos analizę, yra pateiktos (P4 lentelėje, A priedas).

Tiriamai universiteto, kaip valstybinės įstaigos informacinės sistemos, kategorijai yra keliami atitinkami saugumo reikalavimai, kurie remiantis Lietuvos Respublikos vidaus reikalų ministro įsakymu „Dėl valstybės institucijų ir įstaigų informacinių sistemų elektroninės informacijos techninių saugos reikalavimų“ teisės akto Nr. 1V-384 patvirtinto 2008 m. spalio 27 d.“ yra pateikiami B priede.

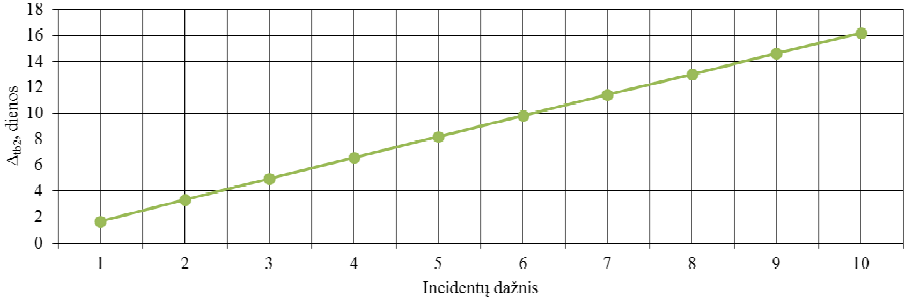
4.3. Informacinės sistemos išliekamumas

Dauguma tiriamų universiteto informacinių sistemų priskiriama trečiajai kategorijai, kur sistemos darbas turi būti atstatytas per 8 val., o informacija pasiekiama 90 % darbo valandomis, kas sudaro 216 valandų per metus. Remiantis P5 lentelės (A priedas) rezultatais buvo nustatyta sistemos modulių atsparumo atakoms charakteristika, kuri esant įgyvendintiems visiems aukščiau aprašytiems apsaugos mechanizmams yra pateikta (P6 lentelėje, A priedas).

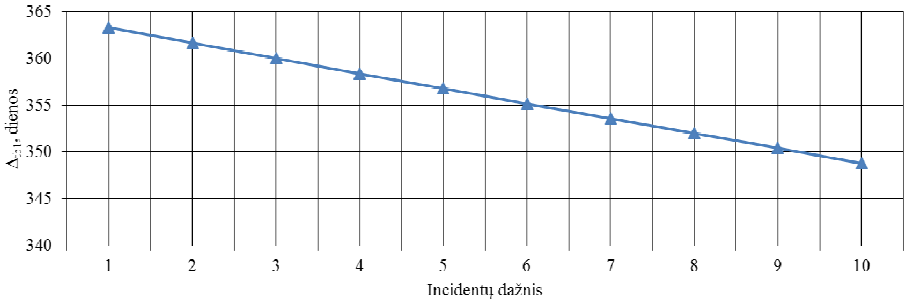
Incidentų sunkumo įtaka modeliuotai sistemai buvo vertinama naudojant skirtingus incidentų rinkinius (nuo vidutinio iki pačio didžiausio sunkumo) remiantis 3.2 lentele. Rezultatams gauti buvo pasirinkta sistema su visais numatomais saugumo mechanizmais. Sistemos laikas sukompromituotoje būsenoje auga didėjant incidentų sunkumui. (P1 pav., a, A priedas), o sistemos būvimas normalioje būsenoje mažėja didėjant incidentų sunkumui (P1 pav., b, A priedas). Incidentų įtaka skirtingoms grėsmėms yra skirtinga dėl informacinių sistemos modulių apsaugos mechanizmų savybių, kurios priklauso nuo sistemos komponentų (P1 pav., c, d, A priedas).

Priklausomai nuo incidentų pasirodymo dažnio modeliuojamos sistemos tikimybė būti sukompromituotai didėja (4.2 pav., a), o sistemos buvimas normalioje būsenoje mažėja didėjant incidentų dažniui (4.2 pav., b). Sistemos būsena, kai yra sukompromituota daugiau negu pusė komponentų, didėjant incidentų pasirodymo dažniui didėja eksponentiškai (4.2 pav., c).

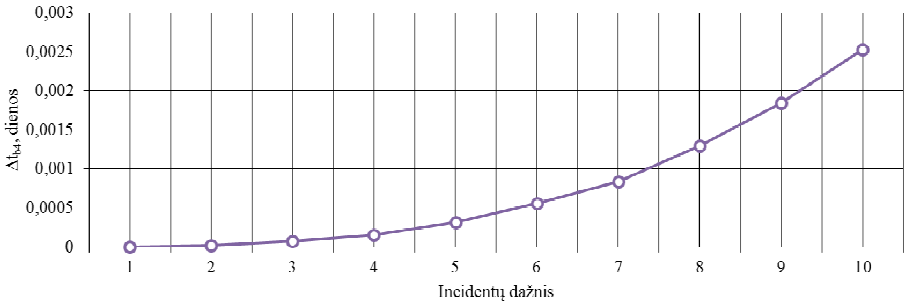
Skirtingi informacinės sistemos moduliai yra apsaugoti skirtingais saugumo mechanizmais ir priklausomai nuo sistemos modulio svorio tikimybė būti sukompromituotam yra skirtinga. Kuo saugumą užtikrinantys mechanizmai yra stipresni, tuo tikimybė sistemai likti normalioje būsenoje yra didesnė (4.3 pav., a), o sukompromituotoje būsenoje tikimybė yra mažiausia (4.3 pav., b).



a)



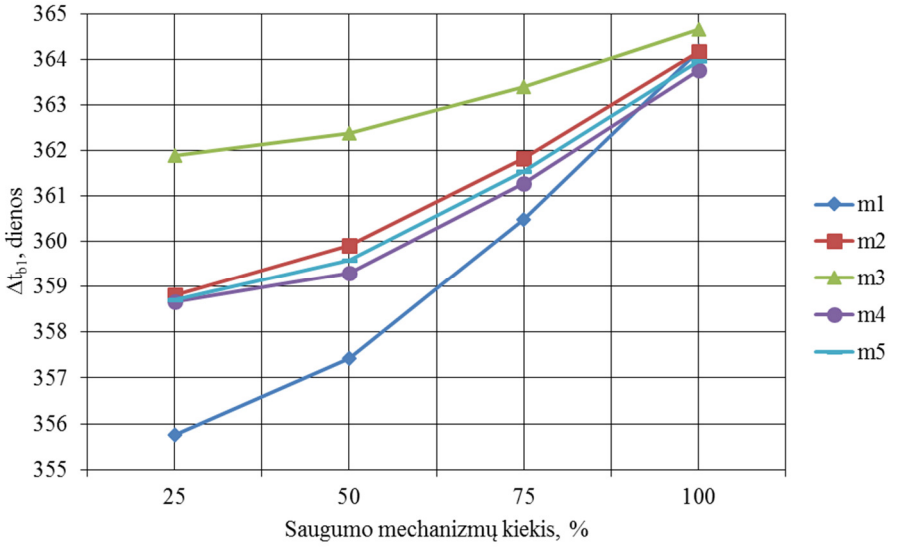
b)



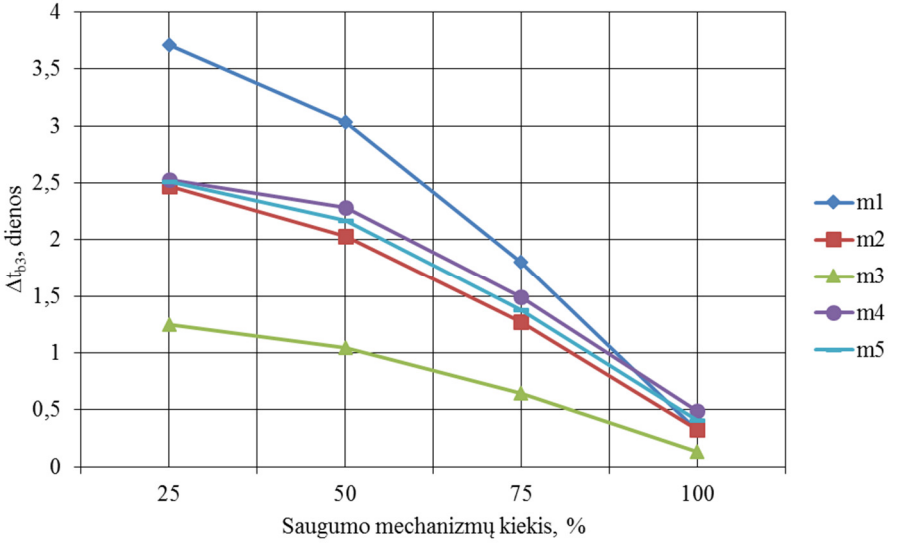
c)

4.2 pav. Incidentų pasirodymo dažnio įtaka sistemos būsenai: a) b_2 b) b_1 ir c) b_4
Fig. 4.2. Incident Occurrence Interval Influence on System States a) b_2 b) b_1 and c) b_4

Saugumo mechanizmų rinkiniai skiriasi tarpusavyje (saugumo mechanizmų kiekis svyruoja nuo 100 % iki 25 %), o jų reikšmė yra išreikšta trečiosios kategorijos informacinės sistemos visų saugumo mechanizmų skaičiumi.



a)

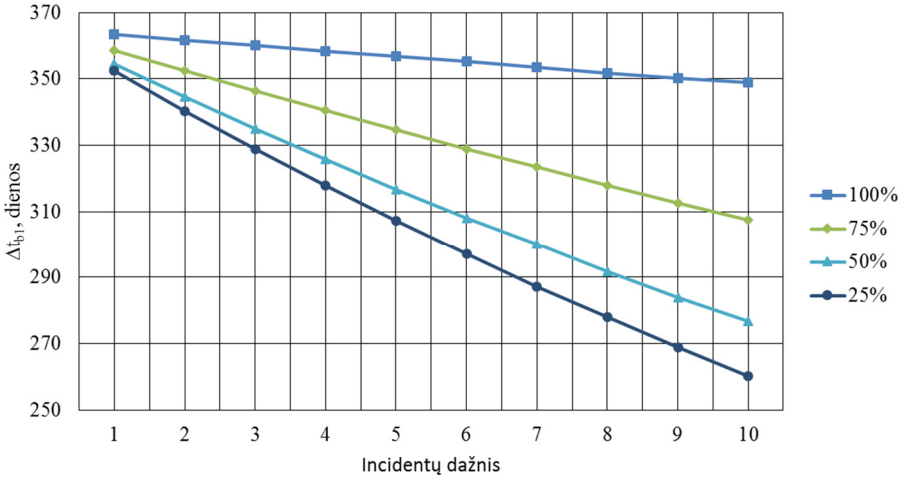


b)

4.3 pav. Saugumo mechanizmų rinkinių įtaka sistemos būsenai (pagal sistemos modulius) ir sistemos atstatymo būsenai a) b_1 ir b) b_3

Fig. 4.3. Protection Mechanism Set Influence on System State according to the Module a) b_1 and b) b_3

Iš 4.3 paveikslo a) ir b) dalies grafikų matome, kad m_3 modulio (elektroninio pašto serveris) atsparumas atakoms mažai priklauso nuo įdiegtų apsaugos mechanizmų kiekio. Modulių m_2 (programinė įranga), m_4 (maršrutizatorius) ir m_5 (ugniasienė) atsparumas atakoms yra beveik vienodas priklausomai nuo apsaugos mechanizmų kiekio, o m_1 modulio (operacinė sistema) atsparumas atakoms labai priklauso nuo įdiegtų saugumo mechanizmų kiekio.



4.4 pav. Saugumo mechanizmų rinkinių įtaka sistemos būsenai pagal incidentų dažnį b_1 sistemos būsena

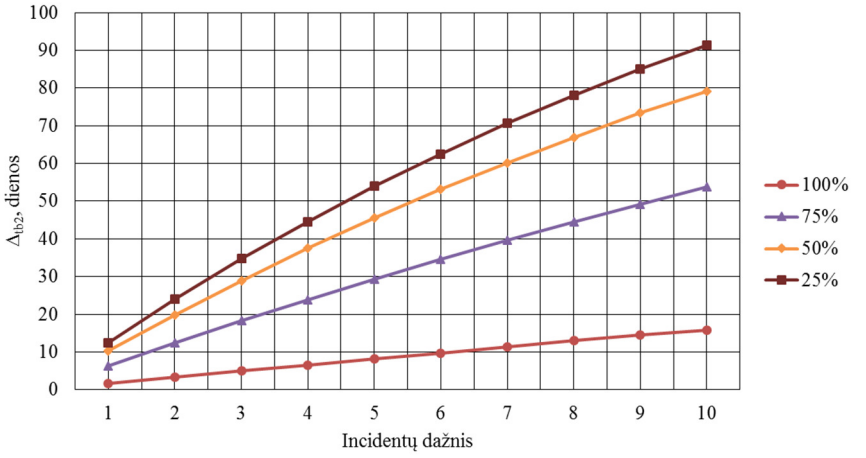
Fig. 4.4. Protection Mechanism Set Influence on System State according to incident frequency b_1 system state

Sistemos būsenų priklausomybė nuo saugumo mechanizmų skaičiaus ir incidentų intensyvumo parodyta grafikuose (4.4 ir 4.5 pav.). Kuo didesnis apsaugos mechanizmų skaičius yra įdiegtas sistemoje, tuo mažiau sistema yra priklausoma nuo incidentų dažnio ir ilgesnį laiką išlieka normalioje būsenoje (4.4 pav.). Kuo didesnis incidentų dažnis ir mažesnis įdiegtų apsaugos mechanizmų skaičius, tuo sistemos laikas sukompromituotoje būsenoje sparčiai didėja (4.5 pav., a). Sistemos būsenos charakteristika, kai pusė sistemos modulių yra sukompromituotų (b_4 būsena) yra pavaizduota 4.5 paveikslo b) dalyje – charakteristikos forma panaši į eksponentę, t. y. didėjant incidentų dažniui sistemos laikas b_4 būsenoje staigiai padidėja.

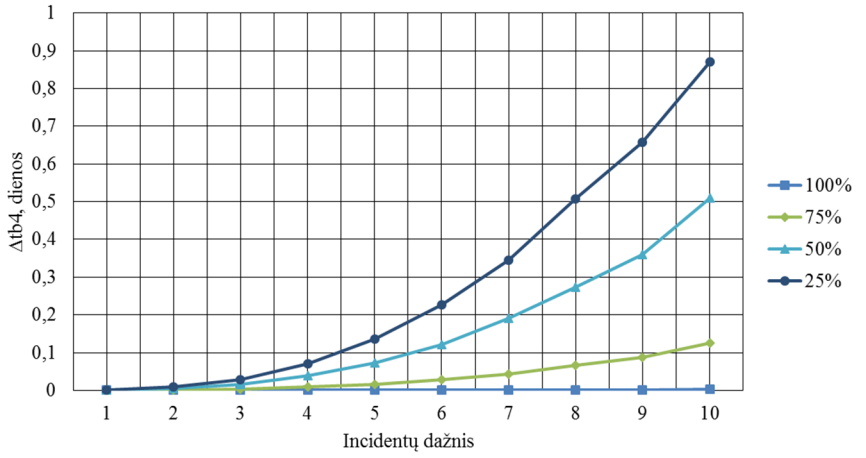
Žemiau yra aprašomos sistemos būsenų b_1, b_2 ir b_4 charakteristikų priklausomybės nuo incidentų dažnio, sistemos saugumo mechanizmų skaičiaus ir sistemos modulio apsaugos mechanizmų nebuvimo.

Charakteristikos buvo gautos keičiant:

- įdiegtų saugumo mechanizmų kiekį nuo 100 % iki 25 %;
- keičiant incidentų pasirodymo dažnį nuo 1 iki 10 kartų;
- eliminuojant modulių saugumo mechanizmus paliekant juos be apsaugos.



a)



b)

4.5 pav. Saugumo mechanizmų rinkinių įtaka sistemos būsenai pagal incidentų dažnį a) b_2 ir b) b_4 sistemos būsenoms

Fig. 4.5. Protection Mechanism Set Influence on System State according to incident frequency a) b_2 and b) b_4 system states

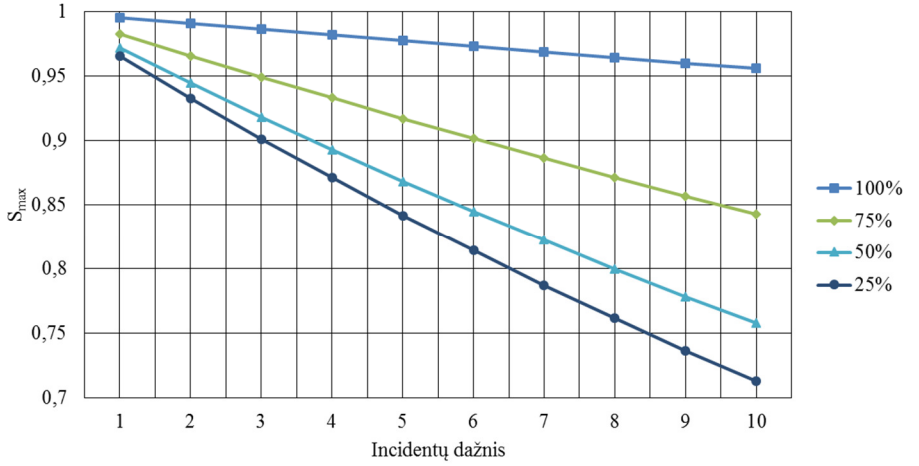
Saugumo mechanizmų rinkinių įtakos sistemos būsenai pagal incidentų dažnį charakteristikos yra pateiktos 4.4 ir 4.5 paveiksle.

Sistemos normalios būsenos b_1, būsenos b_2 ir b_4 priklausomybės nuo incidentų dažnio eliminuojant tam tikrus modulius yra pateiktos P2, P3 ir P4 paveiksluose (A priedas), kur a) 100 %, b) 75 %, c) 50 % ir d) 25 % yra įdiegtų mechanizmų procentinė dalis.

Sistema yra mažiausiai jautri incidentų dažniui, kai joje yra įdiegta 100 % apsaugos mechanizmų, bet ji yra priklausoma nuo modulių apsaugos mechanizmų skaičiaus. Labiausiai priklauso nuo m_1 – operacinės sistemos mechanizmų skaičiaus (nuo 0,6 iki 5,4 % priklausomai nuo incidentų dažnio), mažiausiai nuo m_3 – pašto serverio mechanizmų skaičiaus (nuo 0,3 iki 3,0 % priklausomai nuo incidentų dažnio) ir vienodai priklauso nuo m_2, m_4 ir m_5 modulių apsaugos mechanizmų skaičiaus.

Didžiausią įtaką sistemos saugumui turi m_1 – operacinės sistemos modulio apsaugos mechanizmų kiekis, kai bendroje sistemoje yra įdiegta 75 % visų mechanizmų. Operacinės sistemos modulio įtaka normaliai sistemos būsenai kinta nuo 1,3 iki 9,8 % priklausomai nuo incidentų dažnio.

Mažinant bendrą sistemos apsaugos mechanizmų kiekį iki 50 ir 25 %, visų modulių apsaugos mechanizmų įtaka sistemai tampa beveik lygi ir labai mažai priklauso nuo sistemos apsaugos modulio. Apsaugos mechanizmų įtaka normaliai būsenai kinta nuo 0,1 iki 0,4 %.



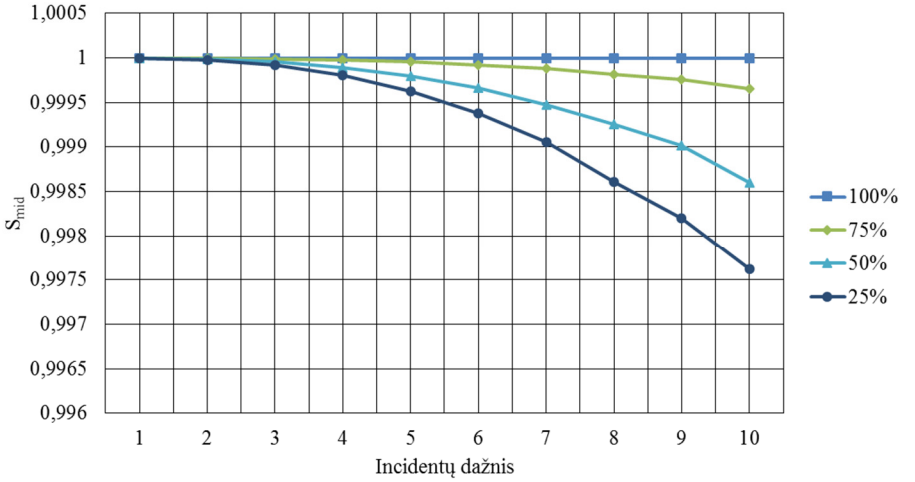
4.6 pav. Saugumo mechanizmų rinkinių įtakos sistemos išliekamumui pagal incidentų dažnį charakteristikos S_{max}

Fig. 4.6. Protection mechanism set influence on system survivability according to incident frequency S_{max}

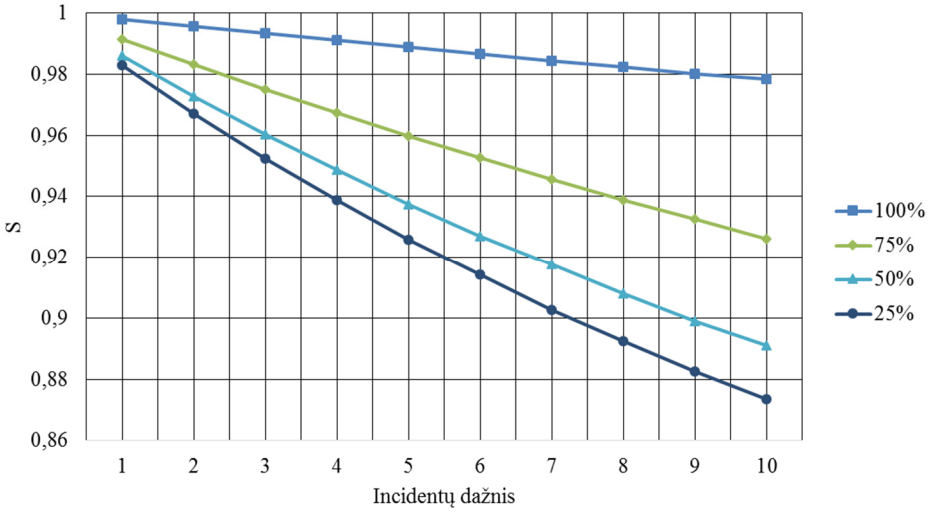
4.3 lentelė. Informacinės sistemos modulių apsaugos charakteristikos
Table 4.3. information system module security charakteristic

Apsaugos mechanizmų kiekis, %	Sistema normalioje būsenoje, dienos		Sistemos pasiekiamumas incidentų dažniui padidėjus x 10
	Incidentų dažnis N x 1	Incidentų dažnis N x 10	
100 %	363,3	348,9	96,0 %
75 %	358,7	307,5	86,0 %
50 %	354,7	276,7	79,6 %
25 %	352,4	260,2	74,7 %

Iš 4.3 lentelės matyti, kad incidentų skaičiui padidėjus 10 kartų, priklausomai nuo įdiegtų apsaugos mechanizmų skaičiaus, sistema sukompromituotoje būsenoje gali būti nuo 14,5 dienų (esant 100 % apsaugos mechanizmų) iki 92 dienų (esant 25 % apsaugos mechanizmų). Projektuojant trečiosios kategorijos informacinės sistemos apsaugos mechanizmų kiekį su atsarga, kad incidentų skaičius padidės 10 kartų, matyti, kad sistemoje įdiegus iki 85 % saugumo mechanizmų sistema tenkins jos kategorijai keliamus reikalavimus, užtikrinančius 90 % sistemos pasiekiamumą darbo dienomis.



4.7 pav. Saugumo mechanizmų rinkinių įtakos sistemos išliekamumui pagal incidentų dažnį charakteristikos S_{mid}
Fig. 4.7. Protection mechanism set influence on system survivability according to incident frequency S_{mid}



4.8 pav. Saugumo mechanizmų rinkinių įtakos sistemos išliekamumui pagal incidentų dažnį charakteristikos *S*
Fig. 4.8. Protection mechanism set influence on system survivability according to incident frequency *S*

Išliekamumas – tai kiekybinė informacinių sistemų saugumo charakteristika, kuri yra parodyta 4.6, 4.7 ir 4.8 paveikslų diagramose. Maksimalus išliekamumas S_{max}, tai tikimybė, kad informacinė sistema po incidento liks normalioje būsenoje. S_{mid} – tai tikimybė, kad pusė informacinės sistemos modulių liks normalioje būsenoje. Informacinės sistemos išliekamumas *S* parodo vidutinę išliekamumo reikšmę, kuri geriausiai atvaizduoja informacinės sistemos saugumo mechanizmų įtaką modeliuojamai sistemai.

4.4. Ketvirtojo skyriaus išvados

1. Skirtingų modulių apsaugos mechanizmų kiekis skirtingai apsaugo nuo incidentų. Mažiausiai nuo saugumo mechanizmų skaičiaus yra priklausomas elektroninio pašto serveris, o daugiausiai – operacinė sistema.

2. Modeliuojamos informacinės sistemos pasiekiamumas labai priklauso nuo incidentų atėjimo dažnio ir apsaugos mechanizmų kiekio. Incidentų dažniui padidėjus 10 kartų, esant 100 % saugumo mechanizmų, sistemos buvimo normalioje būsenoje dienų skaičius sumažėja 14 dienų, o esant 25 % – sumažėja 92 dienomis metų laikotarpyje.

3. Sistema yra mažiausiai jautri incidentų dažniui, kai joje yra įdiegta 100 % apsaugos mechanizmų, bet ji yra priklausoma nuo modulių apsaugos mechanizmų skaičiaus. Labiausiai priklauso nuo operacinės sistemos mechanizmų skaičiaus (nuo 0,6 iki 5,4 % priklausomai nuo incidentų dažnio), mažiausiai nuo pašto serverio mechanizmų skaičiaus (nuo 0,3 iki 3,0 % priklausomai nuo incidentų dažnio) ir vienodai priklauso nuo likusių modulių apsaugos mechanizmų skaičiaus.

4. Projektuojant trečiosios kategorijos informacinės sistemos apsaugos mechanizmų kiekį su atsarga, kad incidentų skaičius padidės 10 kartų, matyti, kad sistemoje įdiegus iki 85 % saugumo mechanizmų, sistema tenkins jos kategorijai keliamus reikalavimus, užtikrinančius 90 % sistemos pasiekiamumą darbo dienomis.

5. Remiantis gauta informacinės sistemos išliekamumo charakteristika galima tiksliai nustatyti koks turi būti įgyvendintas saugumo mechanizmų kiekis sistemoje, kad ji tenkintų keliamus išliekamumo reikalavimus tinkle padidėjus incidentų skaičiui.

Bendrosios išvados

1. Tinklo srauto padalinimas į sekcijas leidžia nustatyti vyraujančias tendencijas tinkle, kurios reikalingos tiksliam tinklo modelio sudarymui. Tai leidžia sumažinti analizuojamų duomenų kiekį ir tiksliau nustatyti paros laikotarpio tinklo srauto tendenciją.

2. Pagal tinklo statistinius duomenis galima modeliuoti tinklo srautą ir nustatyti tinklo normalią būseną. Pagal duomenų paketų atėjimo laiko pasiskirstymo dėsnį galima modeliuoti informacinęs sistemos išliekamumą.

3. Kolmogorov-Smirnov suderinamumo tikrinimo testas parodė, kad TCP ir UDP tinklo srauto paketų atėjimo laiko pasiskirstymo dėsnį tiksliau apibūdinti leidžia Pareto 2 pasiskirstymas negu Puasono. Eksperimentas taip pat parodė, kad TCP ir UDP tinklo srauto pasiskirstymo kreivės yra tokios pat formos.

4. Sukurtas informacinių sistemų išliekamumo statistinis modelis, kuris leidžia įvertinti informacinės sistemos išliekamumą priklausomai nuo incidentų pasiskirstymo, dažnio, svorio, tipo, sistemos modulių svorio, apsaugos mechanizmų skaičiaus, incidentų aptikimo laiko bei modulių atstatymo laiko.

5. Modeliuojamos informacinės sistemos saugumas labiausiai priklauso nuo incidentų pasirodymo tikimybės, apsaugos mechanizmų stiprumo, tuo tarpu incidentų lygis turi mažiausiai įtakos apsaugotai informacinei sistemai.

6. Remiantis gauta informacinės sistemos išliekamumo charakteristika galima tiksliai nustatyti, koks turi būti įgyvendintas saugumo mechanizmų kiekis sistemoje, kad ji tenkintų keliamus išliekamumo reikalavimus tinkle padidėjus incidentų skaičiui.

Literatūra ir šaltiniai

Abry, P.; Veitch, D. 1998. Wavelet analysis of long-range dependent traffic, *IEEE Transactions on Information Theory*, 44:2–15.

Alsubhi, K.; Zhani, M. F.; Boutaba, R. 2012. Embedded Markov Process based Model for Performance Analysis of Intrusion Detection and Prevention Systems, *Global Communications Conference (GLOBECOM)*, IEEE, pp. 898–903.

Annual Report, 2013. *Asia Pacific Computer Emergency Response Team*. [Žiūrėta 2014.08.01] Prieiga per internetą: http://www.apcert.org/documents/pdf/APCERT_Annual_Report_2013(FINAL).pdf

Aussem, A.; Murtagh, F. 2001. Web traffic demand forecasting using wavelet-based multiscale decomposition, *International Journal of Intelligent Systems*, 16:215–236.

Bakhoum, E. 2011. Intrusion detection model based on selective packet sampling, *EURASIP Journal on Information Security.* ISSN 1687-417X.

Barakat, C.; Thiran, P.; Iannaccone, G.; Diot, C.; Owezarski P. 2002. A flow-based model for internet backbone traffic, in *Proc. ACM SIGCOMM Internet Meas.*

Barakat, C.; Thiran, P.; Iannaccone, G.; Diot, C.; Owezarski P. 2003. Modeling internet backbone traffic at the flow level, *IEEE Transactions on Signal Processing.*

Barford P.; Plonka, D. 2001. Characteristics of network traffic flow anomalies. *Proceedings of ACM SIGCOMM Internet Measurement Workshop*, San Francisco, CA.

Barford, P.; Kline, J.; Plonka, D.; Ron, A. 2002. A signal analysis of network traffic anomalies, *Internet Measurement Workshop,* Marseille.

Beaudet, S. T.; Courtney, T.; Sanders, W. H. 2006. A behavior-based process for evaluating availability achievement risk using stochastic activity networks, *Reliability and Maintainability Symposium (RAMS '06)*, pp. 21–28.

Bhattacharjee, A.; Nandi, S. 2010. Statistical analysis of network traffic inter-arrival, in *Proc. of the IEEE 12th International Conference on Advanced Communication Technology (ICACT)*, Feb. 7–10, 2:1052–1057.

Bolch, G.; Greiner, S.; Meer, H.; Trivedi, K. S. 1998. Queueing Networks and Markov Chains, *New York: John Wiley & Sons.*

Cao, J.; Cleveland, W. S.; Gao, Y. 2004. Statistical Models for HTTP Aggregate Source Traffic. Bell Labs, *stat.bell-labs.com, Tech. Rep.*

Cao, J.; Cleveland, W. S.; Lin, D.; Sun, D. X. 2002. Internet Traffic Tends Toward Poisson and Independent as the Load Increases, *Springer*, New York, pp. 83–109.

Cao, J.; Cleveland, W.; Gao, Y.; Jeffay, K.; Smith, F. D.; Weigle, M. 2004. Stochastic Models for Generating Synthetic HTTP Source Traffic, *IEEE INFOCOMM.*

CERT-LT. Lietuvos Respublikos nacionalinis elektroninių ryšių tinklų ir informacijos saugumo incidentų tyrimo padalinys. [Žiūrėta 2014.08.01] Prieiga per internetą: https://www.cert.lt/statistika.html

Clasification Guidelines for Goverment Institution and Office Information Systems according to the Processed Electronic Information 2007, *Order of the Minister of Interior of the Republic of Lithuania. 2007–07–11.* Official Gazette, Nr. 1V–247.

Cleveland, W. S.; Lin, D.; Sun, D. X. 2000. IP Packet Generation: Statistical Models for TCP Start Times Based on Connection-Rate Superposition, *ACM SIGMETRICS*, pp. 166–177.

Crovella, M.; Kolaczyk, E. 2003. Graph wavelets for spatial traffic analysis, *IEEE INFOCOM*, San Francisco.

Deavours, D. D.; Clark, G.; Courtney, T.; Dalys, D.; Derisavi, S.; Doyle, J.M.; Sanders, W. H.; Webster, P. G. 2002. The möbius framework and its implementation, *IEEE Trans. Soft. Eng.* 28:956–969.

DID, 2008. Defence in Depth. Trusted information sharing network for Critical infrastructure protection. [Žiūrėta 2014.08.01] Prieiga per internetą: http://www.tisn.gov.au/Documents/Defence+in+Depth.doc

Elahi, G.; Yu, E.; Zannone, N. 2011. Security Risk Management by Qualitative Vulnerability Analysis, *IEEE Security Measurements and Metrics (Metrisec)*, Third International Workshop, ISBN 978-1-4673-1245-5.

Fras, M.; Mohorko, J.; Cucej, Z. 2008. A new goodness of fit test for histograms regarding network traffic packet size process, in P*roc. of the IEEE International Conference on Advanced Technologies for Communications (ATC),* pp. 345–348.

Garšva, E. 2006a. Computer System Survivability Modelling by Using Stochastic Activity Network. *SAFECOMP'06*, pp. 71–84.

Garšva, E. 2006b. Computer system survivability modelling, *Electronics and Electrical Engineering*, Kaunas: Technologija, 1:48–51.

Garšva, E., Paulauskas N., Gulbinovič L., Stankevičius, A. 2011. Computer System Survivability Evaluation Based on Risk Analysis. Information Systems Architecture and Technology, *Web Information Systems Engineering, Knowledge Discovery and Hybrid Computing Networks*, Wroclaw, pp. 291–301.

Garšva, E.; Paulauskas, N.; Grazulevicius, G.; Gulbinovic, L. 2012. Academic Computer Network Traffic Statistical Analysis, in *Proc. of the IEEE 2nd Baltic Congress on Future Internet Communications (BCFIC)*, pp. 100–105.

Goranin, N.; Cenys A. 2008. Malware propagation modeling by the means of genetic algorithms, *Electronics and Electrical Engineering*, Kaunas: Technologija, 6:23–26.

Goranin, N.; Cenys, A. 2009. Genetic Algorithm Based Internet Worm Propagation Strategy Modeling Under Pressure of Countermeasures, *Journal of Engineering Science and Technology Review*, ISSN 1791–2377, 2:43–47.

Grout, V.; Cunningham, S.; Oram, D.; Hebblewhite, R. 2004. A Note on the Distribution of Packet Arrivals in High-Speed Data Networks, in *Proc. of the IADIS International Conference WWW/Internet 2004*, Oct. 6–9, pp. 889–892.

Gulbinovič, L. 2011. Aštuonių skilčių mikrovaldiklių galimybių tyrimas panaudojimo Ethernet tinklo įrenginiuose, *Mokslas – Lietuvos ateitis. Elektronika ir elektrotechnika*, Vilnius, ISSN 2029-2252, 3:82–86.

Haimes, Y.Y. 2005. Risk modeling, Assessment, and Management, *Harper Collins*, New York, ISBN 978-0-471-72389-9.

Haverkort, B.; Marie, R.; Rubino, G.; Trivedi, K. S. 2001. Performability Modeling Tools and Techniques, *Chichester,* England: John Wiley & Sons.

Heidari, M. 2006. The Role of Modeling and Simulation in Information Security The Lost Ring. Mega Security. [Žiūrėta 2014.08.01] Prieiga per internetą: http://www.megasecurity.org/papers/The_Role_of_Modeling_and_Simulation_in_Information_Security.pdf

Hirel, C.; Tuffin, B.; Trivedi, K. S. 2000. SPNP: Stochastic Petri nets. version 6.0. in Computer Performance Evaluation, *Modelling Techniques and Tools – 11th Int. Conf., TOOLS 2000, ser. Lecture Notes in Computer Science,* Schaumburg, IL, USA: Springer Verlag, 1786:354–357.

Holl, K. 2003. OSI Defense in Depth to Increase Application Security, *Global Information Assurance Certification Paper.* [Žiūrėta 2014.08.01] Prieiga per internetą: http://www.giac.org/paper/gsec/2868/osi-defense-in-depth-increase-application-security/104841

IAIC, 2012. Internet and internet comunications, *The cyber security forum iniciative.* [Žiūrėta 2014.08.01] Prieiga per internetą: http://www.lockheedmartin.com/content/dam/lockheed/data/corporate/documents/LM-White-Paper-Intel-Driven-Defense.pdf

ISO 27001. LST ISO/IEC 27001, Informacinės technologijos. Saugumo metodai. Informacijos saugumo valdymo sistemos. Reikalavimai, Li*etuvos standartizacijos departamentas.*

ISO 27002. LST ISO/IEC 27002, Informacinės technologijos. Saugumo metodai. Informacijos saugumo kontrolės priemonių praktikos nuostatai. Reikalavimai, *Lietuvos standartizacijos departamentas.*

ISO 27005. LST ISO/IEC 27005, Informacijos technologija. Saugumo metodai. Informacijos saugumo rizikos valdymas. Reikalavimai, *Lietuvos standartizacijos departamentas.*

ISO 27033. LST ISO/IEC 27033, Informacinės technologijos. Saugumo metodai. Tinklo saugumas. Tinklo saugumo projektavimo ir diegimo gairės. Reikalavimai, *Lietuvos standartizacijos departamentas.*

ISP, 2006. Information Security Policy - A Development Guide for Large and Small Companies, *SANS Institute InfoSec Reading Room.* [Žiūrėta 2014.08.01] Prieiga per internetą: http://www.sans.org/reading-room/whitepapers/policyissues/information-security-policy-development-guide-large-small-companies_1331

ISP, 2012. Information Security Policy University of Oxford 2012. [Žiūrėta 2014.08.01] Prieiga per internetą: https://www.it.ox.ac.uk/media/global/wwwitservicesoxacuk/sectionimages/security/Information_Security_Policy_2012_07.pdf

Yang, N.; Yu, H.; Sun, H.; Qian, Z. 2010. Quantifying Software Security Based on Stochastic Petri nets, *Journal of Computational Information Systems* 6:3049-3056.

Yegneswaran, V.; Barford, P.; Ullrich, J. 2003. Internet intrusions: global characteristics and prevalence, *ACM SIGMETRICS*, San Diego.

Kajackas, A.; Rainys, R. 2011. Estimation of critical components of internet infrastructure, *Elektronika ir elektrotechnika* 4:35–38.

Kajackas, A; Rainys, R; Aputis, A. 2011 Assessment of cyber attacks influence over internet network, *Elektronika ir elektrotechnika* 7:89–92.

Kbar, G. 2009. Security risk analysis based on probability of system failure, attacks and vulnerabilities, *IEEE Computer Systems and Applications*, ISBN 978-1-4244-3807-5.

Lakhina, A.; Papagiannaki, K.; Crovella, M.; Diot, C.; Kolaczyk, E. D.; Taft, N. 2004. Structural analysis of network flows, *ACM SIGMETRICS*, pp. 61–72.

Lakhina, A.; Papagiannaki, K.; Crovella, M.; Diot, C.; Kolaczyk, E. D.; Taft. N. 2003. Analysis of OD Flows (Raw Data), *Technical Report BUCS-2003-021*, Boston University.

Laurutis, R. 2003. Application of Neural Networks for Data Protection Research, *Electronics and Electrical Engineering* 4:61–64.

Li, Z. C.; Zhang, H.; You, Y.; He, T. 2003. Linuxfow: a high speed backbone measurement facility, *Passive and Active Measurement Workshop (PAM),* California.

Lipiński, Z. 2014. Bezpieczeństwo sieci komputerowych. Uniwersytet Opolski. *Sieci Komputerowe.* [Žiūrėta 2014.08.01] Prieiga per internetą: http://www.math.uni.opole.pl/~zlipinski/skW/SieciKomp-15-BezpieczenstwoSieci.pdf

LITNET CERT. Kompiuterinių incidentų tyrimo LITNET tinkluose tarnyba. [Žiūrėta 2014.08.01] Prieiga per internetą: https://cert.litnet.lt/lt/apie-cert

Marsan, M. A.; Balbo,G.; Conte, G.; Donatelli, S.; Franceschinis, G. 1995. Modelling with Generalized Stochastic Petri Nets, in ser. *Series in parallel computing,* John Wiley and Sons.

Meyer, J. F.; Movaghar, A.; Sanders, W. H. 1985. Stochastic Activity Networks: Structure, Behavior, and Application, *Proc. of the Int. Conf. on Timed Petri Nets,* Torino, Italy, pp. 106–115.

Moitra S.D., Konda S.L. 2000. A Simulation Model for Managing Survivability of Networked Information Systems. [Žiūrėta 2014.08.01] Prieiga per internetą: www.cert.org/research/00tr020.pdf

Moore A.P., Ellison R.J., Linger R.C. 2001. Attack Modeling for Information Security and Survivability. [Žiūrėta 2014.08.01] Prieiga per internetą: www.cert.org/archive/pdf/01tn001.pdf

Mushtaq, S. A.; Rizvi, A. 2005. Statistical analysis and mathematical modeling of network (segment) traffic. *Proc. of the IEEE Symposium on Emerging Technologies*, pp. 246–251.

NCB, 2011. Guidline on Information Security Policy, *National Computer Board,* Issue No. 4 Mauricius. [Žiūrėta 2014.08.01] Prieiga per internetą: http://cert-mu.gov.mu/English/Documents/Guidelines/2011/Guideline%20on%20Information%20Security%20Policy.pdf

NetFlow, 2004. Cisco IOS NetFlow Overview, *ITD Product management.* [Žiūrėta 2014.08.01] Prieiga per internetą: http://www.cisco.com/c/dam/en/us/products/collateral/ios-nx-os-software/ios-netflow/prod_presentation0900aecd80311f57.pdf

NetFlow, 2012. Introduction to Cisco IOS NetFlow, *White Paper.* [Žiūrėta 2014.08.01] Prieiga per internetą: http://www.cisco.com/c/en/us/products/collateral/ios-nx-os-software/ios-netflow/prod_white_paper0900aecd80406232.pdf

Nianhua,Y.; Huiqun, Y.; Zhilin, Q.; Hua, S. 2011. Modeling and quantitatively predicting software security based on stochastic Petri nets, *Mathematical and Computer Modelling.* ISSN 0895-7177.

Nicol, D. M.; Sanders, W. H.; Trivedi, K. S. 2004. Model-Based Evaluation: From Dependability to Security, *IEEE Transaction on dependable and secure computing,* pp. 48–65.

Papagiannaki, K.; Taft, N.; Zhang, Z.; Diot, C. 2003. Long-term forecasting of internet backbone traffic: observations and initial models, *IEEE INFOCOM,* San Francisco.

Paulauskas, N.; Garsva, E. 2008. Attacker Skill Level Distribution Estimation in the System Mean Time-to-Compromise, *1st International Conference on Information Technology*, Gdanks, pp. 463–466.

Paulauskas, N.; Garsva, E.; Skudutis, J. 2009. Network Scan Detection Simulation. *Elektronika ir Elektrotechnika*. ISSN 1392-1215, 2:43–46.

Paxson, V.; Floyd, S. 1995. Wide Area Traffic: The Failure of Poisson Modeling, *IEEE/ACM Trans,* Networking, 3:226–244.

Pinsky, M.; Karlin, S. 2001. *An Introduction to Stochastic Modeling*, 4th edition, ISBN 978-0-12-381416-6.

PTAC, 2011b. Privacy Technical Assistance Center. Data Security: Top Threats to Data Protection. [Žiūrėta 2014.08.01] Prieiga per internetą: http://ptac.ed.gov/sites/default/files/issue-brief-threats-to-your-data.pdf

PTAC, 2011a. Privacy Technical Assistance Center. *Data Security Checklist*. [Žiūrėta 2014.08.01] Prieiga per internetą: http://ptac.ed.gov/sites/default/files/ptac-data-security-checklist.pdf

Pukite, J.; Pukite P. 1998. Markov Modeling for Reliability Analysis, *IEEE Press Series.* ISBN 0-7803-3482-5.

Puniškis, D.; Laurutis, R. 2005. The Use of Neuron Networks for the Performance of Epidemics caused by computer viruses, *Electronics and Electrical Engineering,* 4:28–32.

Puniškis, D.; Laurutis, R. 2007. Behavior Statistic based Neural Net Anti-spam Filters, *Electronics and Electrical Engineering,* 6:35–38.

Puniškis, D.; Laurutis, R.; Dirmeikis, R. 2006. An Artificial Neural Nets for Spam email Recognition, *Electronics and Electrical Engineering,* 5:73–76.

Ramanauskaitė, S. 2010. Modeling of SYN Flooding Attacks, *Jaunųjų mokslininkų darbai,* ISSN 1648-8776. 1:331–335.

Ramanauskaitė, S.; Čenys, A. 2009. DoS atakų modeliavimas stochastiniais metodais, *Jaunųjų mokslininkų darbai,* ISSN 1648-8776, 3:97–101.

Ramanauskaitė, S.; Čenys, A. 2011. Modelling of Central Processing Unit Work Denial of Service, in *Proceedings of the 17th International Conference on Information and Software Technologies*, Kaunas: Technologija, ISSN 2029-0020, pp. 99–104.

Ramanauskaitė, S.; Čenys, A. 2011. Stochastinis TCP SYN atakų modelis, *Mokslas – Lietuvos ateitis*. Vilnius: Technika, ISSN 2029-2341, 3:20–24.

Ramanauskaitė, S.; Čenys, A. 2012. Composite Dos Attack Model, *Mokslas - Lietuvos ateitis: Elektronika ir elektrotechnika,* Vilnius: Technika, ISSN 2029-2341, 4:20–26.

Requirements to Technical Security of Goverment Institution and Office Information Systems, 2008-10-27, Nr. 1V-384.

Requirements to Technical Security of Government Institution and Office Information Systems, *Order of the Minister of Interior of the Republic of Lithuania*, 2008-10-27. Nr. 1V–384. Official Gazette, Nr. 127–4866.

Sallhammar, K.; Helvik, B. E.; Knapskog, S. J. 2005. Incorporating Attacker Behavior in Stochastic Models of Security, in *Proceedings of SAM'05,* Las Vegas, USA.

Sanders, W.H. 2010. Möbius: model-based environment for validation of system reliability, availability, security, and performance. Möbius Manual Version 2.3.1, 2010. [Žiūrėta 2014.08.01] Prieiga per internetą: //www.mobius.illinois.edu/manual/MobiusManual.pdf

Service Name and Transport Protocol Port Number Registry. [Žiūrėta 2014.08.01] Prieiga per internetą: http://www.iana.org/assignments/service-names-port-numbers/service-names-port-numbers.xml

Sommers, J.; Barford, P. 2004. Self-configuring network traffic generation, *IMC '04: Proceedings of the 4th ACM SIGCOMM conference on Internet measurement,* New York, USA: ACM, pp. 68–81.

Surman, G. 2002. Understanding Security Using the OSI Model. *SANS Institute InfoSec Reading Room.* [Žiūrėta 2014.08.01] Prieiga per internetą: http://www.sans.org/reading-room/whitepapers/protocols/understanding-security-osi-model_377

Trivedi, K. S. 2001. Probability and Statistics with Reliability, *Queuing, and Computer Science Applications,* John Wiley and Sons, New York.

Vishwanath, K. V.; Vahdat, A. 2006. Realistic and responsive network traffic generation, in *Proc. of – 2006 Conference on Applications, Technologies, Architectures, and Protocols For Computer Communications,* Pisa.

Zhang, H.; Li, X. 2003. A Scalable High Performance Network Monitoring Agent for CERNET, *PDCAT03.*

Autoriaus mokslinių publikacijų disertacijos tema sąrašas

Straipsniai recenzuojamuose mokslo žurnaluose

Gulbinovič, L. 2011. Aštuonių skilčių mikrovaldiklių panaudojimo Ethernet tinklo įrenginiuose galimybių tyrimas, *Mokslas – Lietuvos Ateitis: Elektronika ir elektrotechnika,* 3:82–86, ISSN 2029-2341. (IndexCopernicus)

Paulauskas, N.; Garšva, E.; Gulbinovič, L.; Stankevičius, A.; Poviliauskas, D. 2012. Survivability Modelling of Lithuanian Government Information System, *Electronics and electrical engineering,* 4:95–98, ISSN 1392-1215. (Science Citation Index Expanded (Web of Science))

Gulbinovič, L. 2012. Kompiuterių sistemų saugumo modeliavimas, *Mokslas – Lietuvos Ateitis: Elektronika ir elektrotechnika,* 4:27–30, ISSN 2029-2341. (IndexCopernicus)

Garšva, E.; Paulauskas, N.; Gražulevičius, G.; Gulbinovič, L. 2014. Packet Inter-arrival Time Distribution in Academic Computer Network, *Electronics and electrical engineering,* 3:87–90, ISSN 1392-1215. (Science Citation Index Expanded (Web of Science))

Straipsniai kituose leidiniuose

Garšva, E.; Paulauskas, N.; Gulbinovič, L.; Stankevičius, A. 2011. Computer System Survivability Evaluation Based on Risk Analysis, *Information systems architecture and*

technology: web information systems engineering, knowledge discovery and hybrid computing, Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej, 2011. ISBN 9788374936309, pp. 291–301.

Paulauskas, N.; Garšva, E.; Gražulevičius, G.; Gulbinovič, L. 2012. Academic Computer Network Traffic Statistical Analysis, *2nd Baltic Congress on Future Internet Communications (BCFIC)*, 25–27 April 2012, Vilnius, Lietuvam, Piscataway: IEEE, 2012. ISBN 9781467316729, pp. 100–105.

Summary in English

Introduction

Problem formulation

Currently during information system design, operation and maintaining, information security and data survivability are considered with highest priority. The requirements for security become more strict for private and government organizations their products and services. These information systems could be starting from cash machine, medical equipment, plane control system and finalizing nuclear reactor control system. Often such i systems are connected via computer network. All system became very depended on computer network security and reliability. This causes increased potential intruders to computer networks interest, because of this the vulnerability of computer system and computer network constantly increasing.

Relevance of the thesis

It is very important to know the information system security components distribution network security incidents and their impact on the security of information systems. Informaion system security model, speeds up the development process of such systems, lower costs and increased security.

Object of research

Object of the research – the information system survivability and assessment tools.

The aim of the thesis

The aim is to create a tool for survivability of information system assessment in design stage, based on the security requirements, evaluating existing security mechanisms and incidents.

Objectives of the thesis

To achieve the objective these challenges are addressed:

1. Analyze the existing network incidents and determine their impact on the security components – the confidentiality, integrity and availability.
2. Analyze the security modeling techniques of information systems.
3. Perform computer network data flows statistical analysis to find the TCP and UDP packets distribution laws.
4. Develop information systems survivability evaluation model.
5. Identify the information system security level characteristics.

Research methodology

The stochastic methods, the probability distribution methods and statistical methods were applicable in this work. The accidental actions were simulated using stochastic activity network models with simulation tool *Mobius*. The network traffic statistical analysis was performed information collected from *NetFlow* protocol.

Scientific novelty of the thesis

During the research the new results of computer science engineering were obtained:

1. Composed information system survivability modeling method, which allows to determine the probabilities of system compromise based on the threat type, severity and security mechanisms.
2. The *NetFlow* protocol was used for network traffic statistics collection in heavy loaded network. The information collected was used to analyze packages distribution. From that set of statistics for TCP and UDP traffic the distribution characteristic were determined.
3. The new formulas proposed, allowing adapt the probability of compromise based on the threat type, severity, and security mechanisms. Formulas can be easily changed and adapted to the real needs of the information systems.

Practical value of research findings

The well-known methods for information security modeling were examined using probability theory, Markov processes, Petri and stochastic activities network. Their advantages and disadvantages were discussed. The modeling techniques suitable for computer network and information system security modeling were a stochastic activity network modeling. Stochastic activity network methods allow to model dynamic systems

where combinatorial methods were unsuitable. In real network is needed to assess the temporal component of the phase transition, which may depend on events. This is best achieved with stochastic activity networks.

Defended statements

1. Pareto's Law 2 can be used for TCP and UDP network traffic packet arrival time distribution describing in academic network.
2. The proposed method can be used for information systems survivability assessment, having only system security requirements and incidents distribution in computer network.
3. Information systems security level can be determined by its survivability characteristics.

The structure ot the thesis

The scientific work consists of the general characteristic of the dissertation, 4 chapters, conclusions, list of literature, list of publications and addenda. The total scope of the dissertation – 107 pages excluding annexes, 41 pictures and 13 tables.

1. Information system security modeling techniques and modeling methods overview

The network security modeling techniques have been compared. Such modeling techniques as probability theory methods, Markov processes, Petri networks and stochastic activity networks were compared. Every method was introduced with his advantages and disadvantages. From discussed methods one of the relevant was chooses stochastic activity network modeling method. Stochastic activity network allow model the dynamic system behavior where probability theory is impropriety. In stochastic activity network incident can occur by many distribution methods when Markov and Petri network methods operate only with exponential distribution. In real system you should evaluate time value when stochastic activity network allow that.

2. Computer network traffic statistical analysis

The statistical analysis results of an academic computer network traffic using the data gathered with *NetFlow* protocol. Results of the statistical analysis are presented in a visual manner which reveals the tendencies of computer network traffic distributions. Computer network modeling or performance evaluation requires knowledge of the computer network characteristic distribution trends. Distribution according to known statistical laws is very applicable during the research, but the uncertainty remains: do the statistical formulas represent the real situation. The main source for getting the information about the network usage is the monitoring systems which use *NetFlow*

protocol and the faculty has the similar one. Information about the network traffic is limited in *NetFlow* and this research attempts to get sufficient information about the network for future models and anomaly detections. The amount of the information and the number of different ways to interpret it is high, so the statistical results of TCP and UDP protocol are presented. The research of TCP and UDP packet inter-arrival time distributions, outliers are considered and statistical distributions are fitted to the experimental curves. The optimal for the addressed network packet inter-arrival time distributions are determined in this work. This is a first and essential step in network traffic modeling and simulation.

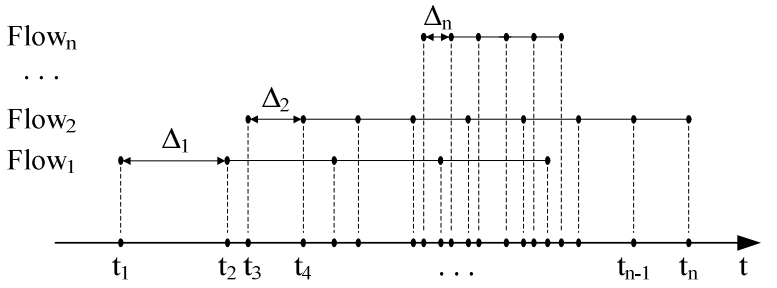


Fig. S2. Assumed distribution of packets in network flows

Table S2. Goodness-of-fit of packet inter-arrival time distribution

Protocol	Distribution	Param. 1	Param. 2	*KS*	*A*	Rank
TCP	Weibull	0.798	0.002	0.063	609.64	2
	Pareto 2	2.658	0.003	0.058	606.03	1
	Gamma	0.509	0.004	0.129	638.19	-
	Exponential	561.680	-	0.123	603.03	-
	Lognormal	1.409	-7.151	0.075	602.14	3
UDP	Weibull	0.734	0.003	0.053	322.35	-
	Pareto 2	1.848	0.004	0.046	322.90	1
	Gamma	0.480	0.008	0.103	333.79	-
	Exponential	249.330	-	0.176	317.91	-
	Lognormal	1.545	-6.472	0.053	323.57	2

Network time characteristics are very important in the modeling. Distribution of time interval between the packets was calculated using the logic presented in Figure S2. Let us assume that all packets in any network flow are distributed with the uniform time intervals from each other, i. e. in Flow1 time interval between packets is equal to Δ_1, in Flow2 – Δ_2 and in Flown – Δ_n. However in real world, packets may be distributed in unequal time intervals between each other, but we can't say about this from collected *NetFlow*. Also we need to pay attention to the fact that several network flows can occur at the same time, i. e. they overlap. As we know, the packets in network are transmitted one after another. After sorting of the arrival times of packets t_1, t_2... t_n from each flow, we obtain that time intervals between packets are different and randomly distributed.

This is explained by the fact that the flows are also randomly distributed in time. From the packet distribution in time, we can calculate the time intervals between packets.

Obtained distribution of time interval between the packets is presented in Figure S3.

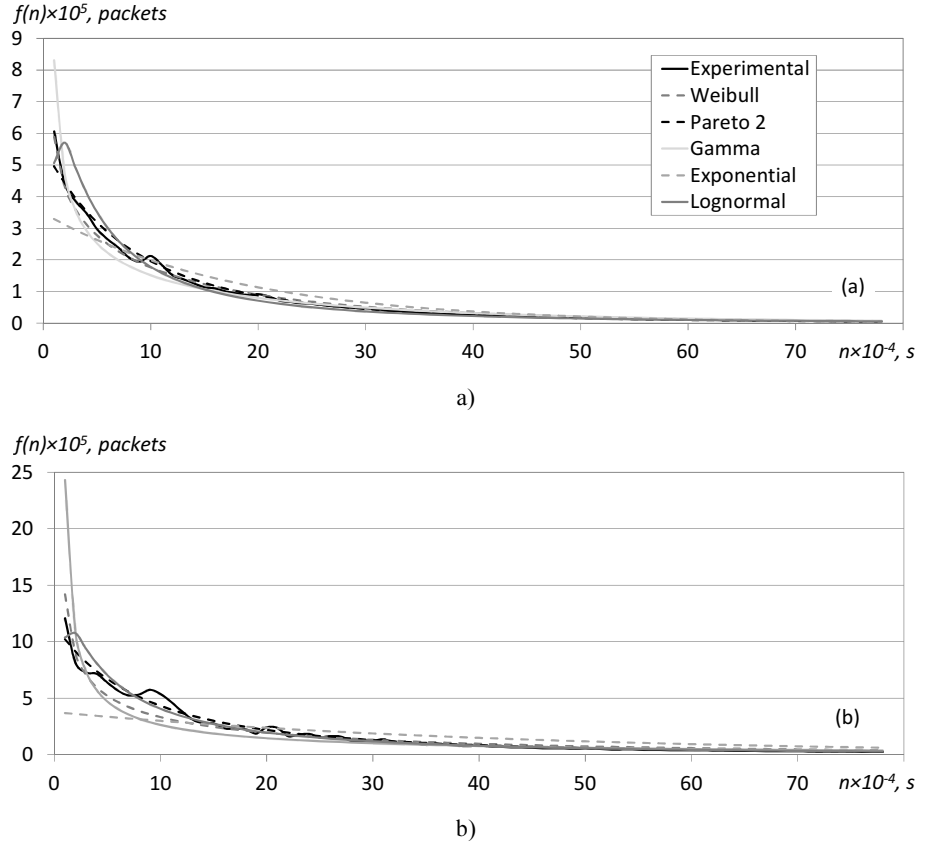


Fig. S4. Goodness-of-fit of packet inter-arrival time distributions: (a) TCP, (b) UDP

Distribution of time interval between the packets shows that packet arrival time has the same form despite their direction. Arrival of the incoming packets is more intense. The network of the faculty is not overloaded, so the biggest part of the packets arrives during the shortest intervals between them.

Best fitting distributions were determined using Kolmogorov-Smirnov goodness-of-fit test. Weibull, Pareto, Gamma, Exponential and Lognormal distributions were considered as they are close to an average experimental distribution curve and are used in computer network traffic modelling. The goodness-of-fit tests were performed for packet inter-arrival time distributions based on the average network traffic going both directions (Table S2). Weibull, Pareto, Lognormal and Gamma distributions use shape

parameter and scale parameter. Exponential distribution uses rate parameter. A shape parameter α affects the shape of a distribution and scale parameter β stretches or shrinks it. The first parameter of distribution in Table V is represented by Param. 1 and the second by Param. 2. Kolmogorov-Smirnov parameter KS shows the maximum absolute difference between the experimental and distribution curves and the lower it is the better is the fit. Multiplier A is used to adjust the fitted distribution to the experimental curve on y axis and it changes the value of Probability Density Function integral to the value of A.

For the average network traffic Pareto Second Kind (Pareto 2) distribution fits best for both protocols as it is seen in Fig. S4.

Pareto Second Kind distribution is a standard Pareto distribution with shifted x axis so it falls into $0 \leq x < +\infty$, while in standard Pareto distribution $\beta \leq x < +\infty$. Pareto Second Kind probability distribution conditionally can be called Lomax distribution and is a heavy-tail distribution usually used in business or economical modelling.

Network traffic sections are considered and goodness-off-fit was performed for all the sections in order to determine its distributions and parameters. CDFs are considered up to 0.99, as other values contribute only to the tails.

All distribution curves for TCP were successfully fitted; fitting UDP curves was challenging (Table S3): 1 and 5 curves were not fitted; best values were presented in the table.

Table S3. Packet inter-arrival time distribution coefficients

Protocol	Section	Distribution	α	β	KS	A
TCP	1	Pareto2	2.7278	0.0034	0.0478	359.63
	2	Weibull	0.8978	0.0011	0.0623	1890.70
	3	Pareto2	1.9981	0.0022	0.0558	698.54
	4	Pareto2	3.1931	0.0361	0.0246	23.18
	5	Weibull	0.7750	0.0027	0.0441	230.12
	6	Gamma	0.7810	0.0023	0.0545	1201.46
	7	Weibull	0.7578	0.0028	0.0467	437.86
	8	Weibull	0.7897	0.014	0.0211	24.93
UDP	1	Lognormal	1.9497	-6.084	0.0560	46.55
	2	Pareto2	1.8484	0.0039	0.0456	322.91
	3	Pareto2	1.2351	0.0009	0.0643	339.50
	4	Pareto2	4.0406	0.1065	0.0221	10.35
	5	Weibull	0.6066	0.0073	0.0476	40.96
	6	Pareto2	1.9606	0.0042	0.0409	338.45
	7	Pareto2	1.4612	0.0022	0.0454	184.25
	8	Pareto2	2.7942	0.0386	0.0293	10.15

Pareto 2 fits most of the curves, especially those where the traffic is growing or falling. Pareto 2 distribution was the best or second best for all the distributions except TCP section 6, so Pareto 2 distribution can be chosen to model network packet inter-arrival time.

3. Information system survivability evaluation according risk analysis

Information system survivability simulation model was composed and information system survivability simulation was performed in this research. Model parameters were taken from legal regulation and risk analysis. Requirements to the system recovery time and accessibility are set by regulation, based on the information system category. The simulation was done by using stochastic activity networks. Simulation results show that the modelled information system security mostly depends on the incident occurrence probability, on the strength of protection mechanisms, while the occurring incident severity has the least effect on the protected information system.

The knowledge about the system and threats occurring to it is vital, in order to evaluate information systems security. This information can be revealed by the risk analysis. Survivability is a common, numeric characteristic of system ability to survive the incident. It is used for system comparison, and security mechanism evaluation. During this research information system survivability simulation model was composed. Model parameters were taken from information system security regulation and risk analysis. Modelling results are presented and analyzed. The model is different from the previous model, because of a slightly different information system concept.

Government information systems are best regulated and will be addressed in this research. Information systems are categorized based on its vitality to the state. Then, the requirements to the system recovery time and accessibility are set. There are four different categories. The requirements for 1st and 2nd system categories are very high and system recovery time should not be longer when 15 min for first category and one hour for second category. The information accessibility is set as 99% for 1st and 96% for 2nd system category. The requirements for 3rd and 4th system categories are set only for working hours and working days. We chose to model the system which is qualified as to be in third category, because most governmental systems are managed by different institutions in Lithuania and they are not directly regulated by the law. Third category information system must be recovered in 8 hours and must be accessible 90% at working hours. Security mechanisms are described for each category. These requirements are summarized in Table S4. Adapting these regulations to the modelled environment we consider that there are 36 different mechanisms, one distinct mechanism can be used to protect from one or more threats.

Table S4. Requirements to the Lithuanian government system accessibility and recovery time

Category	System recovery time	Information accessibility	Number of subsystems
I	15 min	99%	7
II	1 h	96%	5
III	8 wh	90% wd	3
IV	16 wh	70% wd	2

The regulation sets the information system complexity requirements: 4th category must have 2 or more information system subsystems (modules), 3rd category must have no less than 3 modules, 2nd category no less than 5 and 1st category no less than 7 modules. We made an assumption that modelled information system despite being in third category is complex and has 5 modules m.

The requirements for first and second system categories are very high and system recovery time should not be longer when 15 minutes for first category and 1 hour for second category. The information accessibility is set as 99 percent for first and 96 percent for second system category. The requirements for third and fourth system categories are set only for working hours and working days. Also each category has no less than a specified number of subsystems. Requirements to main security mechanisms, which must be implemented in each category, are different. Every upper category system must have additional security mechanisms alongside security mechanisms which are specified for lower category systems.

According to requirements which are presented in the Table 1 the four categories models were composed. The creation of models and simulation results are presented in the following sections.

Risk analysis in this research addresses one aspect of all information security – the computer network risks rising from the outer perimeter of the computer system. During the risk analysis assets of the hypothetical third category information system were identified, threats were outlined and implemented security mechanisms revised. There is different amount of security mechanisms implemented to protect different modules in the modelled information system.

Risk analysis showed module compromise detection interval Δtd, the importance of different modules described as module weight *w(m)* and the rate of different information system module usage, so module usage probabilities *PM(m)* were determined. Then the probabilities of incidents targeted to confidentiality (*PCm(j)*), integrity (*PIm(j)*) and availability (*PAm(j)*) for different modules according to incident severities *Pm(j)* were determined.

The information system is a distributed computer network with boundaries defined, which is facing the computer incident i after the time interval $\Delta t_{inc,i}$ and withstands it or one or more modules of the system are compromised on time *tc,i*. The degradation of the system is detected after some time $\Delta t_{d,i}$ and then the system state is restored after the interval of time $\Delta t_{r,i}$. Information system is modelled during the time interval Δt_{all}, which is long enough for all the events to appear (Fig. S5).

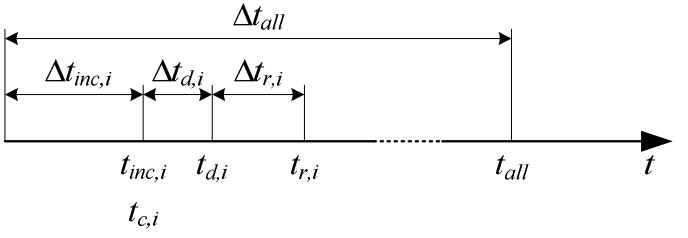


Fig. S5. Information system security events

Incident occurrence in this model is a stochastic process, not a genetic algorithm which is often used to model malware propagation. Incidents are grouped according to the threat and can have different severity levels j, first one is most severe $j = 1$. Incidents occur by Poisson law and target the specific module m of the modelled information system considering its rate of use. All the system modules are protected by the security mechanisms Na defined by the risk analysis. Information system modules have different importance which is represented by its weigh *w(m)*. Structure of information system model is presented in Fig. S6. Model parameter values were determined by risk analysis or set by regulation; corresponding parts are marked on the models structure.

After the incident information systems module is compromised or not, by that affecting the state of the whole system. The information system can be in five states: normal (b1), when no modules are compromised; one module compromised (b2); when one or more modules are compromised and the system state change is detected and the system goes to the recovery state (b3); if the incident is severe then more than half of the system modules can be compromised (b4); or even all of them can be damaged (b5).

It's more likely that the systems degradation will be detected faster and the problem will be addressed in the shorter interval than the one set by the law, that's why triangular distribution is used.

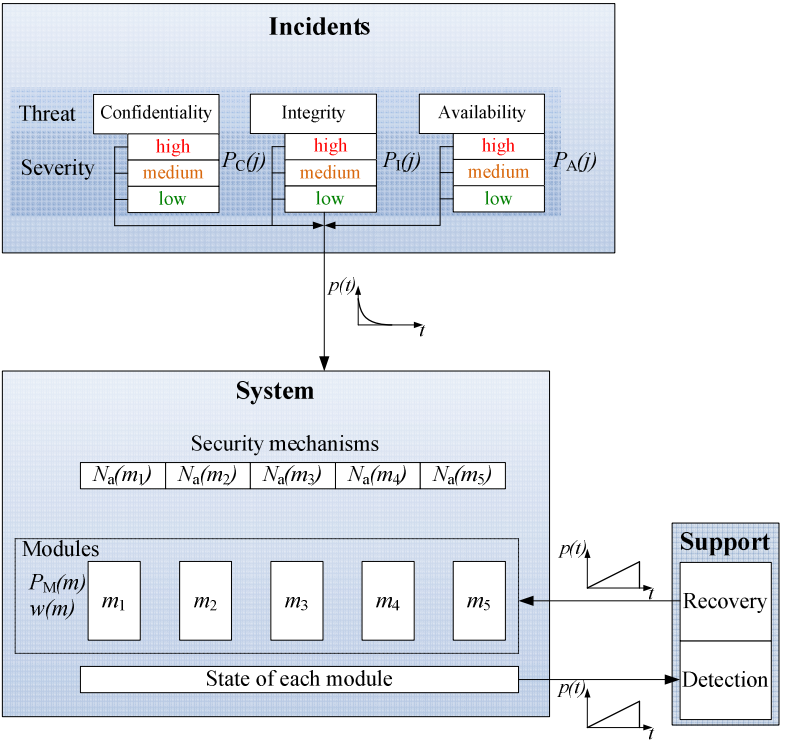


Fig. S6. Information system model

Information system simulation model is composed using Stochastic Activity Network (SAN) formalism, which is quite similar to the stochastic Petri nets. Simulation model organized using Mobius tool by its design repeats the block diagram of the model.

Information system survivability S is universal and quantitative characteristic showing its ability to provide the services it is intended to in the hostile environment, which influences the level of the provided service.

When service or the system survives in the maximal functional state b_1 during the system usage time Δt_{all} then such characteristic can be called maximal survivability S_{max}:

$$S_{max} = \frac{\Delta t_{b1}}{\Delta t_{all}}. \tag{S2}$$

If system survives in the functional state, which represent half of its functionality described by functional state b_4 during the system usage time (Δt_{all}) then such characteristic can be called midrange survivability S_{mid}:

$$S_{mid} = \frac{\Delta t_{b4}}{\Delta t_{all}}. \tag{S3}$$

Different services or modules providing these services represent different importance to the mission of the system, this must be considered. Survivability of the system S can be describeed as:

$$S = \sum_m w(m)S(m),\ 0 \le S(m) \le 1,\ \sum w(m) = 1,\ 0 \le w(m) \le 1. \tag{S4}$$

Where $S(m)$ is the survivability of information system module m, and $w(m)$ is the weight of the module.

4. Academic information system security evaluation

In this chapter are discussed the reliability of information systems components. The real network of information systems, security incidents and security requirements analysis that generated the statistical distributions have been adapted to information system security modeling. By taking advantage of the distribution received by the statistics obtained by the academic network of informaion systems security level evaluation of the performance of the conclusions and the security of the system-level test.

Security incident – real or potential adverse effects on a information system or computer network activity an event whose outcome – deception, loss, abuse, threat information, ownership of the loss or damage to it. For example, the penetration of information systems, technical exploits, computer viruses or other unwanted software installation. Three key concepts relevant to information security in computer networks is confidentiality, integrity and availability. The concepts associated with the people who use the information resources of identity, the powers and checks liability. Once the information is scanned or copied the result is called a loss of confidentiality. When transferred to the network information is replaced by an unforeseen way, it is called a

loss of integrity. When information is deleted and/or becomes unavailable, it is called a loss of reachability information. The case of a legitimate user is unable to access the network or its individual services, it is experiencing denial of service.

The analysis of each type of incident can create a table for each attack incident system security components. Table S5 contains information about the incident impact on the confidentiality, integrity and availability.

In order to get on the network security incidents affect the confidentiality, integrity and availability of the distribution is required for computer network safety equipment that captures the amount of security incidents and determine their type. Most systems are suitable for use in Intrusion Detection Systems. The data was collected in Vilnius Gediminas Technical University, Faculty of Electronic Intrusion Detection System.

Table S5. Security incident influence on confidentiality, integrity and availability

Security incidents	Influence on system		
	Confidentiality	Integrity	Availability
DoS		+	+
Malware	+	+	+
Trojan Horse	+	+	
Spam		+	+
Port Scaning	+	+	
Phishing	+		
System Compromise / Intrusion	+	+	+
Spyware	+		+
Social Engineering	+	+	+

Intrusion Detection Systems data have been collected in 2012 for all 12 months of reported incidents. The report includes information about such incidents, how to recognize viruses, *Spam* attack and denial of service attacks.

The names of the viruses have been requested for a detailed description of the virus and it was determined by a specific effect of the virus system of the confidentiality, integrity and accessibility and the level of attack. The virus type was assumed that viruses affect the confidentiality, integrity and availability of the same and the attack rate also distributes evenly.

Regardless of the *Spam* sender and the contents of this type of attack is attributed to attacks that affect the integrity and availability, and the level of attacks attributed to the medium.

According to denial of service attacks type the attack has been requested for a detailed description, and was determined by a specific influence on the system confidentiality, integrity and accessibility and the level of attack. The type of attack has been assumed that the entire system is affected by security and is equally confidentiality, integrity and availability, and the attack rate is also divided equally.

In intrusion detection system the flooding attack packets are counted as separate events. Because of the large number of events viruses incidents represent only 0,005% and 0.185% of all spam incidents. Due to the adoption of certain conditions and flooding, spamming incidents normalization. That the number of events to be displayed in a solid number of attacks will be accepted subject to flooding, sync / sessions and scanning attacks, respectively composed of 10 000, 1 000 and 100 packages. Spam incident number was divided by 10. Such conditions were taken based on real flood attack flows. Because of spam there was assumed that for every 100 sent a letter reaches the addressee mailbox. After the implementation of the conditions, it was revealed that 29.8% of viruses, spam messages are 11.0% and 59.2% flooding attack all attacks.

According to a security incident analysis was obtained by the probability distribution of incidents by severity of incidents: 3 – the lightest 1 – the most difficult, and in accordance with the incident impact the confidentiality, integrity and availability. Incident distribution values are shown in Table S6.

Table S6. Incident distribution

Treat	Incident severity		
	1	2	3
Confidentiality	0.148	0.052	0.061
Integrity	0.120	0.084	0.132
Availability	0.293	0.054	0.055

The test of a information system is composed of five modules such as: *m1* – the operating system, *m2* – application, *m3* – e-mail server, *m4* – router *m5* – firewall. Modules shows the importance weight *w(m)* and the different security modules use demonstrates the use of the frequency probability $P_{M(m)}$. Incidents that are distributed according to the corresponding probabilities: confidentiality ($P_{CM\ (j)}$), integrity ($P_{IM\ (j)}$) and reach ($P_{am(j)}$) corresponding to different modules and threats difficulties $P_{m(j)}$.

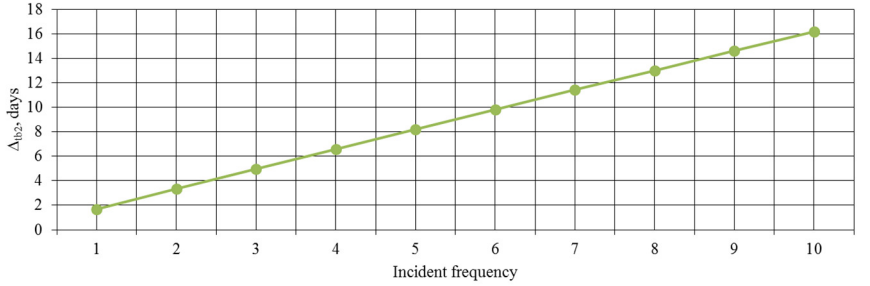
According to the safety requirements the set for each information system module resistance of confidentiality, integrity and availability attacks were presented. Security mechanisms have been attributed to the information system modules.

Our studied university system most are assigned to the third category of systems. The third category of information systems must be restored within 8 hours and information is available 90% of the hours representing 216 hours per year.

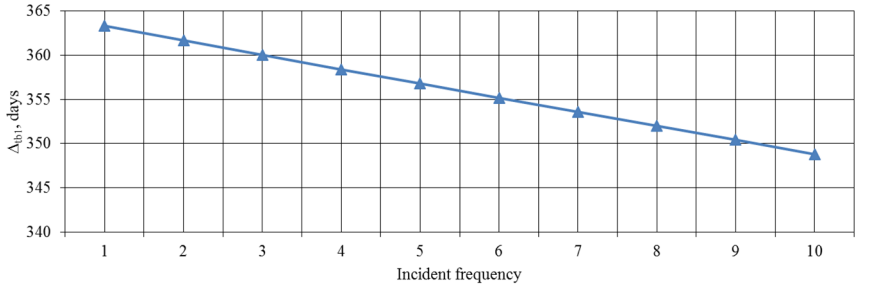
Depending on the incident frequency performance of the modeled system increases the probability of being compromised (Figure S10, a), and the existence of a system in a normal state decreases with increasing frequency of incidents (Figure S10, b). State of the system is compromised when more than half of the components increases the frequency of incidents is increasing exponentially appearance (Figure S10, c).

The different modules of the information system are protected the security mechanisms and different depending on the system module weight probability to be compromised is different. The mechanisms to ensure security is stronger the higher the probability to stay in the normal state of the system is increased (Figure S11, a), and the compromised state is least likely (Figure S11, b).

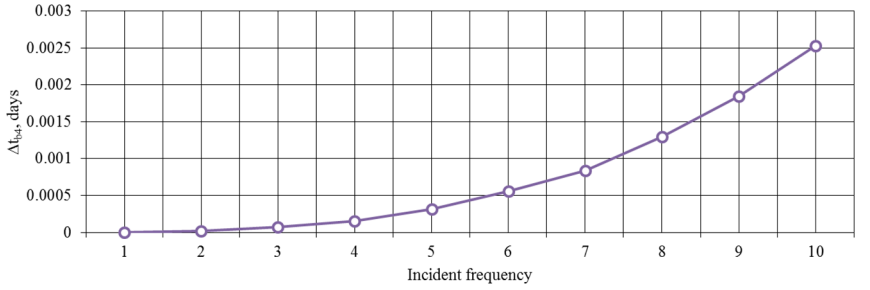
Different security mechanisms sets from 100% to 25% of the security mechanisms of points. Security mechanisms for different sets with each other and their value are expressed as a safety mechanism to the third category of information system security mechanisms numbers.



a)



b)

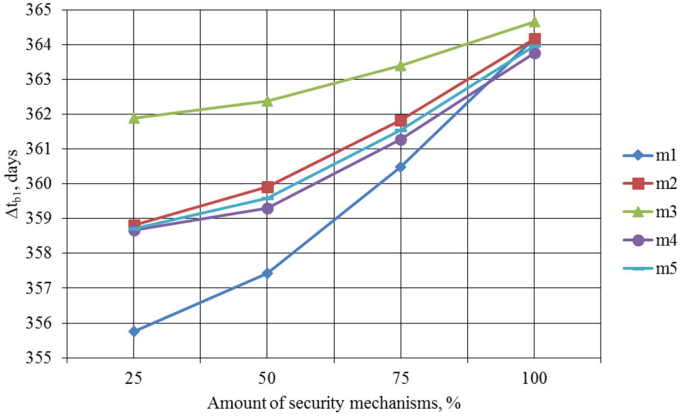


c)

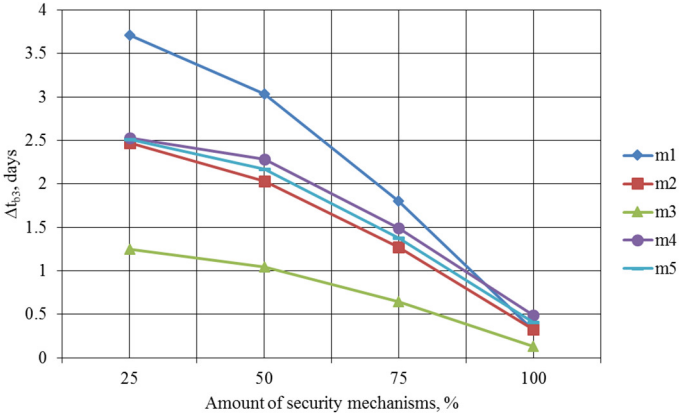
Fig. S10. Incident Occurrence Interval Influence on System States a) *b2* b) *b1* and c) *b4*

In (Figure S11, a, b) we see that m_3 (email server) module have little resistance to attack depends on installed content protection mechanisms. m_2 modules (applications), m_4 (router) and m_5 (firewall) resistance to attacks is almost the same, depending on the

content protection mechanisms, and m_1 (operating system) module resistance attacks is highly dependent on the amount of tan security mechanisms.



a)



b)

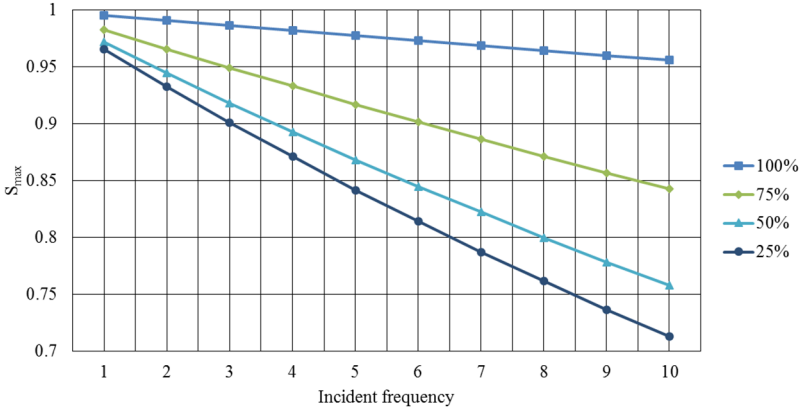
Fig. S11. Protection Mechanism Set Influence on System State according to the Module a) *b1* and b) *b3*

State of the system dependencies on the security mechanisms and the number of incident intensity is shown in (Figure S11, a, b, c) diagrams. If in system is installed more number of security mechanisms, the system is less dependent on the frequency of incidents and the longer term is in normal mode (Figure S11, a). For higher frequency of incidents and less implement security mechanisms number the system compromised state is rapidly increasing (Figure S11, b). System state characterization of the system modules are half b_4 compromise state is shown (Figure S11, c) the characteristics of a

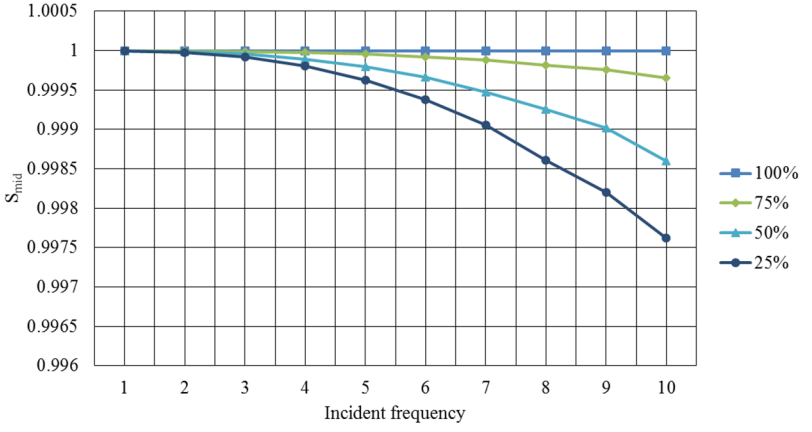
shape similar to the exponent. Increase in frequency of incidents of system time b_4 state sharply increased.

Below are the system states b_1, b_2 and b_4 characteristics dependence on the frequency of incidents, system security mechanisms, and the number of the module system protection mechanisms exist. Characteristics were obtained by changing:

- implement security mechanisms in an amount from 100% to 25%;
- changing the appearance of the frequency of incidents from 1 to 10 times;
- eliminating modules security mechanisms leaving them without protection.



a)



b)

Fig. S12. Protection mechanism set influence on system survivability according to incident frequency a) S_{max}, b) S_{mid}

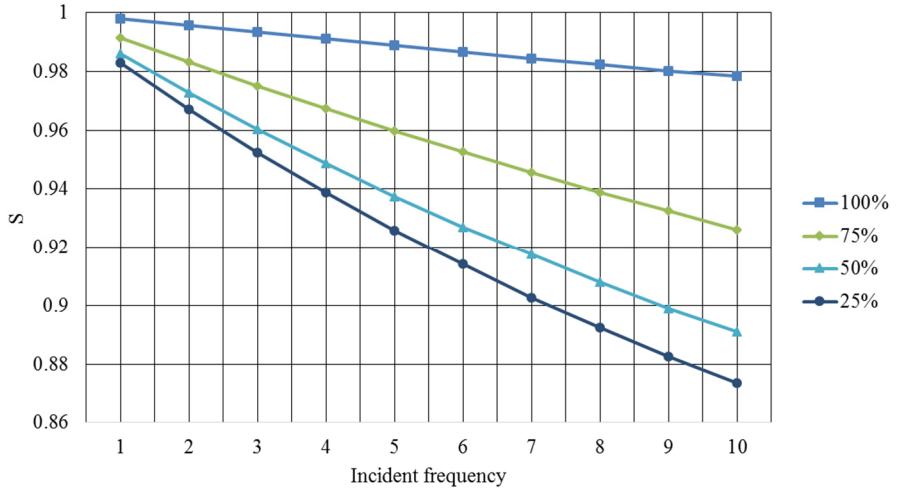


Fig. S13. Protection mechanism set influence on system survivability according to incident frequency *S*

The system is the least sensitive to the frequency of incidents, when it is introduced with 100% protection mechanisms, but is dependent on the number of modules in protection mechanisms. Most depend on m_1 – the operating system using the number (from 0.6 to 5.4 %, depending on the frequency of incidents), the lowest since m_3 - mail server number of machines (from 0.3 to 3.0%, depending on the frequency of incidents) and the same depends on m_2, m_4 and m_5 modules number of protection mechanisms.

The greatest impact on the system security has m_1 – the operating system of the module content protection mechanisms, the overall system is installed in 75% of all machines. Operating system module influence the normal state of the system varies from 1.3 to 9.8%, depending on the frequency of incidents.

By reducing the overall system security mechanisms amount to 50, 25% of all modules of security mechanisms, the influence of the system becomes almost equal and very little dependent on the system for the module. A security mechanism for the influence of the normal state varies from 0,1 to 0.4%.

In Table S7, we see that in the number of incidents increased 10 times, depending on the number of security mechanisms the system compromised state can be from 14.5 days at 100% protection mechanisms for up to 92 days at 25% of the security mechanisms. The design of the third category of information system security mechanisms, the amount of the reserve, the number of incidents increased 10 times, we see that the introduction of the system to 85% of the security mechanisms of the system will satisfy the category requirements to ensure system availability of 90% on weekdays.

Table S7. Information system module security characteristic

Security mechanism set, %	System in normal state, days		System survivability increased incident by 10 times
	Incident frequency, N x1	Incident frequency, N x10	
100%	363.3	348.9	96.0%
75%	358.7	307.5	86.0%
50%	354.7	276.7	79.6%
25%	352.4	260.2	74.7%

Survivability is a quantitative characteristic of information system security, which is shown in (Figure S12, a, b and S13) diagrams. Maximum survivability S_{max} is the probability that the information system after the incident remain normal state. Smid is a chance that half of the information system modules remain in the normal state. Information system survivability S shows the average importance of survivability is the best representative of the information system security mechanisms influence the system model.

General conclusions

1. Network traffic division into sections is reasonable, as it reveals dominant trends which are needed in order to compose general network model and to choose definitive statistical distribution.
2. Network statistics can be used for modeling of network traffic, can also be used to determine the normal state of the network. During information system survivability modeling the results and inter-arrival time distribution results were used.
3. Kolmogorov-Smirnov goodness-of-fit test shows that Pareto Second Kind distribution fits best for both TCP and UDP network packet inter-arrival time distribution experimental curves, this also show that TCP and UDP network packet inter-arrival time distributions are of the same shape.
4. Information system survivability simulation model was designed and implemented. The model allows evaluating information system survivability according to incident severity, occurrence frequency, threat category (CIA) and protection mechanism strength.
5. The modelled information system security mostly depends on the incident occurrence probability, on the strength of protection mechanisms, while the occurring incident severity has the least effect on the protected information system.
6. Based on the information system survivability characteristics, we can accurately determine what has to be implemented with security mechanisms for system to satisfy the requirements for information sysytem survivability by increase of incidents.

Priedai[1]

A priedas. Informacinės sistemos saugumo incidentai. Lentelės ir grafikai
B priedas. Informacinės sistemos saugumo reikalavimai
C priedas. Bendraautorių sutikimai teikti publikacijų medžiagą disertacijoje
D priedas. Autoriaus mokslinių publikacijų disertacijos tema kopijos

[1] Priedai pateikiami pridėtoje kompaktinėje plokštėje

Lech GULBINOVIČ

INFORMACINIŲ SISTEMŲ SAUGUMO TYRIMAS IR IŠLIEKAMUMO VERTINIMO MODELIO SUKŪRIMAS

Daktaro disertacija

Technologijos mokslai,
Informatikos inžinerija (07T)

INFORMATION SYSTEM SECURITY EVALUATION AND CREATION OF SURVIVABILITY EVALUATION MODEL

Doctoral Dissertation

Technological Sciences,
Informatics engineering (07T)

2014 11 14. 10,0 sp. l. Tiražas 20 egz.
Vilniaus Gedimino technikos universiteto
leidykla „Technika“,
Saulėtekio al. 11, 10223 Vilnius,
http://leidykla.vgtu.lt
Spausdino UAB „Ciklonas“
J. Jasinskio g. 15, 01111 Vilnius